(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization International Bureau



(43) International Publication Date 9 August 2001 (09.08.2001)

PCT

(10) International Publication Number WO 01/57188 A2

(51) International Patent Classification7:

_ _ _

Jose, CA 95117 (US). **DRMANAC, Radoje, T.** [YU/US]; 850 East Greenwich Place, Palo Alto, CA 94303 (US).

Glovsky, and Popeo, P.C., One Financial Center, Boston,

- (21) International Application Number: PCT/US01/03800
- (22) International Filing Date: 5 February 2001 (05.02.2001)
- (25) Filing Language:

English

C12N

(26) Publication Language:

English

(30) Priority Data:

09/496,914 3 February 2000 (03.02.2000) US 09/560,875 27 April 2000 (27.04.2000) US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US 09/496,914 (CIP)
Filed on 3 February 2000 (03.02.2000)
US 09/560,875 (CIP)
Filed on 27 April 2000 (27.04.2000)

- (71) Applicant (for all designated States except US): HYSEQ, INC. [US/US]; 670 Almanor Avenue, Sunnyvalle, CA 94086 (US).
- (72) Inventors; and
- (75) Inventors/Applicants (for US only): TANG, Y., Tom [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). LIU, Chenghua [CN/US]; 1125 Ranchero Way, #14, San

850 East Greenwich Place, Palo Alto, CA 94303 (US).

(74) Agent: ELRIFI, Ivor, R.; Mintz, Levin, Coh, Ferris,

MA 02111 (US).

- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

 without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

NOVEL NUCLEIC ACIDS AND POLYPEPTIDES

1. TECHNICAL FIELD

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with uses for these polynucleotides and proteins, for example in therapeutic, diagnostic and research methods.

2. BACKGROUND

5

10

15

20

25

30

35

Technology aimed at the discovery of protein factors (including *e.g.*, cytokines, such as lymphokines, interferons, CSFs, chemokines, and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (*i.e.*, partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization-based cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity, for example, by virtue of their secreted nature in the case of leader sequence cloning, by virtue of their cell or tissue source in the case of PCR-based techniques, or by virtue of structural similarity to other genes of known biological activity.

Identified polynucleotide and polypeptide sequences have numerous applications in, for example, diagnostics, forensics, gene mapping; identification of mutations responsible for genetic disorders or other traits, to assess biodiversity, and to produce many other types of data and products dependent on DNA and amino acid sequences.

3. SUMMARY OF THE INVENTION

The compositions of the present invention include novel isolated polypeptides, novel isolated polynucleotides encoding such polypeptides, including recombinant DNA molecules, cloned genes or degenerate variants thereof, especially naturally occurring variants such as allelic variants, antisense polynucleotide molecules, and antibodies that specifically recognize one or more epitopes present on such polypeptides, as well as hybridomas producing such antibodies.

The compositions of the present invention additionally include vectors, including expression vectors, containing the polynucleotides of the invention, cells genetically engineered to contain such polynucleotides and cells genetically engineered to express such polynucleotides.

The present invention relates to a collection or library of at least one novel nucleic acid sequence assembled from expressed sequence tags (ESTs) isolated mainly by sequencing by hybridization (SBH), and in some cases, sequences obtained from one or more public databases. The invention relates also to the proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins. These nucleic acid sequences are designated as SEQ ID NO: 1-1350. The polypeptides sequences are designated SEQ ID NO: 1351-2700. The nucleic acids and polypeptides are provided in the Sequence Listing. In the nucleic acids provided in the Sequence Listing, A is adenosine; C is cytosine; G is guanine; T is thymine; and N is any of the four bases. In the amino acids provided in the Sequence Listing, * corresponds to the stop codon.

5

10

15

20

25

30

The nucleic acid sequences of the present invention also include, nucleic acid sequences that hybridize to the complement of SEQ ID NO:1-1350 under stringent hybridization conditions; nucleic acid sequences which are allelic variants or species homologues of any of the nucleic acid sequences recited above, or nucleic acid sequences that encode a peptide comprising a specific domain or truncation of the peptides encoded by SEQ ID NO:1-1350. A polynucleotide comprising a nucleotide sequence having at least 90% identity to an identifying sequence of SEQ ID NO:1-1350 or a degenerate variant or fragment thereof. The identifying sequence can be 100 base pairs in length.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of SEQ ID NO:1-1350.

A collection as used in this application can be a collection of only one polynucleotide. The collection of sequence information or identifying information of each sequence can be provided on a nucleic acid array. In one embodiment, segments of sequence information is provided on a nucleic acid array to detect the polynucleotide that contains the segment. The array can be designed to detect full-match or mismatch to the polynucleotide that contains the segment. The collection can also be provided in a computer-readable format.

This invention also includes the reverse or direct complement of any of the nucleic acid sequences recited above; cloning or expression vectors containing the nucleic acid sequences; and host cells or organisms transformed with these expression vectors. Nucleic acid sequences (or their reverse or direct complements) according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology, such as use as hybridization probes, use as primers for PCR, use in an array, use in computer-readable media, use in sequencing

full-length genes, use for chromosome and gene mapping, use in the recombinant production of protein, and use in the generation of anti-sense DNA or RNA, their chemical analogs and the like.

In a preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids of the invention are used as primers in expression assays that are well known in the art. In a particularly preferred embodiment, the nucleic acid sequences of SEQ ID NO:1-1350 or novel segments or parts of the nucleic acids provided herein are used in diagnostics for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

10

20

25

The isolated polynucleotides of the invention include, but are not limited to, a polynucleotide comprising any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; a polynucleotide comprising any of the full length protein coding sequences of SEQ ID NO:1 - 1350; and a polynucleotide comprising any of the nucleotide sequences of the mature protein coding sequences of SEQ ID NO: 1-1350. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent hybridization conditions to (a) the complement of any one of the nucleotide sequences set forth in SEQ ID NO:1-1350; (b) a nucleotide sequence encoding any one of the amino acid sequences set forth in the Sequence Listing (e.g., SEQ ID NO: 1351-2700); (c) a polynucleotide which is an allelic variant of any polynucleotides recited above; (d) a polynucleotide which encodes a species homolog (e.g. orthologs) of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of any of the polypeptides comprising an amino acid sequence set forth in the Sequence Listing.

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising any of the amino acid sequences set forth in the Sequence Listing; or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides with biological activity that are encoded by (a) any of the polynucleotides having a nucleotide sequence set forth in SEQ ID NO:1-1350; or (b) polynucleotides that hybridize to the complement of the polynucleotides of (a) under stringent hybridization conditions. Biologically or immunologically active variants of any of the polypeptide sequences in the Sequence Listing, and "substantial equivalents" thereof (e.g., with at least about 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% or 99% amino acid sequence identity) that preferably retain biological activity are also contemplated. The polypeptides of the invention may be wholly or partially chemically synthesized but are preferably produced by recombinant means using the genetically engineered cells (e.g. host cells) of the invention.

The invention also provides compositions comprising a polypeptide of the invention. Polypeptide compositions of the invention may further comprise an acceptable carrier, such as a hydrophilic, *e.g.*, pharmaceutically acceptable, carrier.

The invention also provides host cells transformed or transfected with a polynucleotide of the invention.

5

10

15

20

25

30

35

The invention also relates to methods for producing a polypeptide of the invention comprising growing a culture of the host cells of the invention in a suitable culture medium under conditions permitting expression of the desired polypeptide, and purifying the polypeptide from the culture or from the host cells. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Polynucleotides according to the invention have numerous applications in a variety of techniques known to those skilled in the art of molecular biology. These techniques include use as hybridization probes, use as oligomers, or primers, for PCR, use for chromosome and gene mapping, use in the recombinant production of protein, and use in generation of anti-sense DNA or RNA, their chemical analogs and the like. For example, when the expression of an mRNA is largely restricted to a particular cell or tissue type, polynucleotides of the invention can be used as hybridization probes to detect the presence of the particular cell or tissue mRNA in a sample using, e.g., in situ hybridization.

In other exemplary embodiments, the polynucleotides are used in diagnostics as expressed sequence tags for identifying expressed genes or, as well known in the art and exemplified by Vollrath et al., Science 258:52-59 (1992), as expressed sequence tags for physical mapping of the human genome.

The polypeptides according to the invention can be used in a variety of conventional procedures and methods that are currently applied to other proteins. For example, a polypeptide of the invention can be used to generate an antibody that specifically binds the polypeptide. Such antibodies, particularly monoclonal antibodies, are useful for detecting or quantitating the polypeptide in tissue. The polypeptides of the invention can also be used as molecular weight markers, and as a food supplement.

Methods are also provided for preventing, treating, or ameliorating a medical condition which comprises the step of administering to a mammalian subject a therapeutically effective amount of a composition comprising a polypeptide of the present invention and a pharmaceutically acceptable carrier.

In particular, the polypeptides and polynucleotides of the invention can be utilized, for example, in methods for the prevention and/or treatment of disorders involving aberrant protein expression or biological activity.

The present invention further relates to methods for detecting the presence of the polynucleotides or polypeptides of the invention in a sample. Such methods can, for example, be utilized as part of prognostic and diagnostic evaluation of disorders as recited herein and for the identification of subjects exhibiting a predisposition to such conditions. The invention provides a method for detecting the polynucleotides of the invention in a sample, comprising contacting the sample with a compound that binds to and forms a complex with the polynucleotide of interest for a period sufficient to form the complex and under conditions sufficient to form a complex and detecting the complex such that if a complex is detected, the polynucleotide of interest is detected. The invention also provides a method for detecting the polypeptides of the invention in a sample comprising contacting the sample with a compound that binds to and form a complex with the polypeptide under conditions and for a period sufficient to form the complex and detecting the formation of the complex such that if a complex is formed, the polypeptide is detected.

10

15

20

25

30

35

The invention also provides kits comprising polynucleotide probes and/or monoclonal antibodies, and optionally quantitative standards, for carrying out methods of the invention. Furthermore, the invention provides methods for evaluating the efficacy of drugs, and monitoring the progress of patients, involved in clinical trials for the treatment of disorders as recited above.

The invention also provides methods for the identification of compounds that modulate (i.e., increase or decrease) the expression or activity of the polynucleotides and/or polypeptides of the invention. Such methods can be utilized, for example, for the identification of compound that can ameliorate symptoms of disorders as recited herein. Such methods can include, but are not limited to, assays for identifying compounds and other substances that interact with (e.g., bind to) the polypeptides of the invention. The invention provides a method for identifying a compound that binds to the polypeptides of the invention comprising contacting the compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and detecting the complex by detecting the reporter gene sequence expression such that if expression of the reporter gene is detected the compound the binds to a polypeptide of the invention is identified.

The methods of the invention also provides methods for treatment which involve the administration of the polynucleotides or polypeptides of the invention to individuals exhibiting symptoms or tendencies. In addition, the invention encompasses methods for treating diseases o disorders as recited herein comprising administering compounds and other substances that modulate the overall activity of the target gene products. Compounds and other substances can

effect such modulation either on the level of target gene/protein expression or target protein activity.

The polypeptides of the present invention and the polynucleotides encoding them are also useful for the same functions known to one of skill in the art as the polypeptides and polynucleotides to which they have homology (set forth in Table 2). If no homology is set forth for a sequence, then the polypeptides and polynucleotides of the present invention are useful for a variety of applications, as described herein, including use in arrays for detection.

4. DETAILED DESCRIPTION OF THE INVENTION

4.1 DEFINITIONS

10

15

20

25

30

35

It must be noted that as used herein and in the appended claims, the singular forms "a", "an" and "the" include plural references unless the context clearly dictates otherwise.

The term "active" refers to those forms of the polypeptide which retain the biologic and/or immunologic activities of any naturally occurring polypeptide. According to the invention, the terms "biologically active" or "biological activity" refer to a protein or peptide having structural, regulatory or biochemical functions of a naturally occurring molecule. Likewise "immunologically active" or "immunological activity" refers to the capability of the natural, recombinant or synthetic polypeptide to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

The term "activated cells" as used in this application are those cells which are engaged in extracellular or intracellular membrane trafficking, including the export of secretory or enzymatic molecules as part of a normal or disease process.

The terms "complementary" or "complementarity" refer to the natural binding of polynucleotides by base pairing. For example, the sequence 5'-AGT-3' binds to the complementary sequence 3'-TCA-5'. Complementarity between two single-stranded molecules may be "partial" such that only some of the nucleic acids bind or it may be "complete" such that total complementarity exists between the single stranded molecules. The degree of complementarity between the nucleic acid strands has significant effects on the efficiency and strength of the hybridization between the nucleic acid strands.

The term "embryonic stem cells (ES)" refers to a cell that can give rise to many differentiated cell types in an embryo or an adult, including the germ cells. The term "germ line stem cells (GSCs)" refers to stem cells derived from primordial stem cells that provide a steady and continuous source of germ cells for the production of gametes. The term "primordial germ

cells (PGCs)" refers to a small population of cells set aside from other cell lineages particularly from the yolk sac, mesenteries, or gonadal ridges during embryogenesis that have the potential to differentiate into germ cells and other cells. PGCs are the source from which GSCs and ES cells are derived The PGCs, the GSCs and the ES cells are capable of self-renewal. Thus these cells not only populate the germ line and give rise to a plurality of terminally differentiated cells that comprise the adult specialized organs, but are able to regenerate themselves.

The term "expression modulating fragment," EMF, means a series of nucleotides which modulates the expression of an operably linked ORF or another EMF.

As used herein, a sequence is said to "modulate the expression of an operably linked sequence" when the expression of the sequence is altered by the presence of the EMF. EMFs include, but are not limited to, promoters, and promoter modulating sequences (inducible elements). One class of EMFs are nucleic acid fragments which induce the expression of an operably linked ORF in response to a specific regulatory factor or physiological event.

10

15

20

25

30

35

The terms "nucleotide sequence" or "nucleic acid" or "polynucleotide" or "oligonculeotide" are used interchangeably and refer to a heteropolymer of nucleotides or the sequence of these nucleotides. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA) or to any DNA-like or RNA-like material. In the sequences herein A is adenine, C is cytosine, T is thymine, G is guanine and N is A, C, G or T (U). It is contemplated that where the polynucleotide is RNA, the T (thymine) in the sequences provided herein is substituted with U (uracil). Generally, nucleic acid segments provided by this invention may be assembled from fragments of the genome and short oligonucleotide linkers, or from a series of oligonucleotides, or from individual nucleotides, to provide a synthetic nucleic acid which is capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon, or a eukaryotic gene.

The terms "oligonucleotide fragment" or a "polynucleotide fragment", "portion," or "segment" or "probe" or "primer" are used interchangeably and refer to a sequence of nucleotide residues which are at least about 5 nucleotides, more preferably at least about 7 nucleotides. more preferably at least about 9 nucleotides, more preferably at least about 11 nucleotides and most preferably at least about 17 nucleotides. The fragment is preferably less than about 500 nucleotides, preferably less than about 200 nucleotides, more preferably less than about 100 nucleotides, more preferably less than about 50 nucleotides and most preferably less than 30 nucleotides. Preferably the probe is from about 6 nucleotides to about 200 nucleotides. preferably from about 15 to about 50 nucleotides, more preferably from about 17 to 30 nucleotides and most preferably from about 20 to 25 nucleotides. Preferably the fragments can

be used in polymerase chain reaction (PCR), various hybridization procedures or microarray procedures to identify or amplify identical or related parts of mRNA or DNA molecules. A fragment or segment may uniquely identify each polynucleotide sequence of the present invention. Preferably the fragment comprises a sequence substantially similar to any one of SEQ ID NOs:1-1350.

5

10

15

20

25

30

Probes may, for example, be used to determine whether specific mRNA molecules are present in a cell or tissue or to isolate similar nucleic acid sequences from chromosomal DNA as described by Walsh et al. (Walsh, P.S. et al., 1992, PCR Methods Appl 1:241-250). They may be labeled by nick translation, Klenow fill-in reaction, PCR, or other methods well known in the art. Probes of the present invention, their preparation and/or labeling are elaborated in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY; or Ausubel, F.M. et al., 1989, Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, both of which are incorporated herein by reference in their entirety.

The nucleic acid sequences of the present invention also include the sequence information from the nucleic acid sequences of SEQ ID NO:1-1350. The sequence information can be a segment of any one of SEQ ID NO:1-1350 that uniquely identifies or represents the sequence information of that sequence of SEQ ID NO:1-1350. One such segment can be a twenty-mer nucleic acid sequence because the probability that a twenty-mer is fully matched in the human genome is 1 in 300. In the human genome, there are three billion base pairs in one set of chromosomes. Because 4²⁰ possible twenty-mers exist, there are 300 times more twenty-mers than there are base pairs in a set of human chromosomes. Using the same analysis, the probability for a seventeen-mer to be fully matched in the human genome is approximately 1 in 5. When these segments are used in arrays for expression studies, fifteen-mer segments can be used. The probability that the fifteen-mer is fully matched in the expressed sequences is also approximately one in five because expressed sequences comprise less than approximately 5% of the entire genome sequence.

Similarly, when using sequence information for detecting a single mismatch, a segment can be a twenty-five mer. The probability that the twenty-five mer would appear in a human genome with a single mismatch is calculated by multiplying the probability for a full match $(1 \div 4^{25})$ times the increased probability for mismatch at each nucleotide position (3×25) . The probability that an eighteen mer with a single mismatch can be detected in an array for expression studies is approximately one in five. The probability that a twenty-mer with a single mismatch can be detected in a human genome is approximately one in five.

The term "open reading frame," ORF, means a series of nucleotide triplets coding for amino acids without any termination codons and is a sequence translatable into protein.

The terms "operably linked" or "operably associated" refer to functionally related nucleic acid sequences. For example, a promoter is operably associated or operably linked with a coding sequence if the promoter controls the transcription of the coding sequence. While operably linked nucleic acid sequences can be contiguous and in the same reading frame, certain genetic elements e.g. repressor genes are not contiguously linked to the coding sequence but still control transcription/translation of the coding sequence.

The term "pluripotent" refers to the capability of a cell to differentiate into a number of differentiated cell types that are present in an adult organism. A pluripotent cell is restricted in its differentiation capability in comparison to a totipotent cell.

10

15

20

25

The terms "polypeptide" or "peptide" or "amino acid sequence" refer to an oligopeptide, peptide, polypeptide or protein sequence or fragment thereof and to naturally occurring or synthetic molecules. A polypeptide "fragment," "portion," or "segment" is a stretch of amino acid residues of at least about 5 amino acids, preferably at least about 7 amino acids, more preferably at least about 9 amino acids and most preferably at least about 17 or more amino acids. The peptide preferably is not greater than about 200 amino acids, more preferably less than 150 amino acids and most preferably less than 100 amino acids. Preferably the peptide is from about 5 to about 200 amino acids. To be active, any polypeptide must have sufficient length to display biological and/or immunological activity.

The term "naturally occurring polypeptide" refers to polypeptides produced by cells that have not been genetically engineered and specifically contemplates various polypeptides arising from post-translational modifications of the polypeptide including, but not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation and acylation.

The term "translated protein coding portion" means a sequence which encodes for the full length protein which may include any leader sequence or any processing sequence.

The term "mature protein coding sequence" means a sequence which encodes a peptide or protein without a signal or leader sequence. The "mature protein portion" means that portion of the protein which does not include a signal or leader sequence. The peptide may have been produced by processing in the cell which removes any leader/signal sequence. The mature protein portion may or may not include the initial methionine residue. The methionine residue may be removed from the protein during processing in the cell. The peptide may be produced synthetically or the protein may have been produced using a polynucleotide only encoding for the mature protein coding sequence.

The term "derivative" refers to polypeptides chemically modified by such techniques as ubiquitination, labeling (e.g., with radionuclides or various enzymes), covalent polymer attachment such as pegylation (derivatization with polyethylene glycol) and insertion or substitution by chemical synthesis of amino acids such as omithine, which do not normally occur in human proteins.

5

10

15

20

25

30

35

The term "variant" (or "analog") refers to any polypeptide differing from naturally occurring polypeptides by amino acid insertions, deletions, and substitutions, created using, e.g., recombinant DNA techniques. Guidance in determining which amino acid residues may be replaced, added or deleted without abolishing activities of interest, may be found by comparing the sequence of the particular polypeptide with that of homologous peptides and minimizing the number of amino acid sequence changes made in regions of high homology (conserved regions) or by replacing amino acids with consensus sequence.

Alternatively, recombinant variants encoding these same or similar polypeptides may be synthesized or selected by making use of the "redundancy" in the genetic code. Various codon substitutions, such as the silent changes which produce various restriction sites, may be introduced to optimize cloning into a plasmid or viral vector or expression in a particular prokaryotic or eukaryotic system. Mutations in the polynucleotide sequence may be reflected in the polypeptide or domains of other peptides added to the polypeptide to modify the properties of any part of the polypeptide, to change characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate.

Preferably, amino acid "substitutions" are the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, *i.e.*, conservative amino acid replacements. "Conservative" amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues involved. For example, nonpolar (hydrophobic) amino acids include alanine, leucine, isoleucine, valine, proline, phenylalanine, tryptophan, and methionine; polar neutral amino acids include glycine, serine, threonine, cysteine, tyrosine, asparagine, and glutamine; positively charged (basic) amino acids include arginine, lysine, and histidine; and negatively charged (acidic) amino acids include aspartic acid and glutamic acid. "Insertions" or "deletions" are preferably in the range of about 1 to 20 amino acids, more preferably 1 to 10 amino acids. The variation allowed may be experimentally determined by systematically making insertions, deletions, or substitutions of amino acids in a polypeptide molecule using recombinant DNA techniques and assaying the resulting recombinant variants for activity.

Alternatively, where alteration of function is desired, insertions, deletions or non-conservative alterations can be engineered to produce altered polypeptides. Such alterations

can, for example, alter one or more of the biological functions or biochemical characteristics of the polypeptides of the invention. For example, such alterations may change polypeptide characteristics such as ligand-binding affinities, interchain affinities, or degradation/turnover rate. Further, such alterations can be selected so as to generate polypeptides that are better suited for expression, scale up and the like in the host cells chosen for expression. For example, cysteine residues can be deleted or substituted with another amino acid residue in order to eliminate disulfide bridges.

5

10

15

20

25

30

35

The terms "purified" or "substantially purified" as used herein denotes that the indicated nucleic acid or polypeptide is present in the substantial absence of other biological macromolecules, e.g., polynucleotides, proteins, and the like. In one embodiment, the polynucleotide or polypeptide is purified such that it constitutes at least 95% by weight, more preferably at least 99% by weight, of the indicated biological macromolecules present (but water, buffers, and other small molecules, especially molecules having a molecular weight of less than 1000 daltons, can be present).

The term "isolated" as used herein refers to a nucleic acid or polypeptide separated from at least one other component (e.g., nucleic acid or polypeptide) present with the nucleic acid or polypeptide in its natural source. In one embodiment, the nucleic acid or polypeptide is found in the presence of (if anything) only a solvent, buffer, ion, or other component normally present in a solution of the same. The terms "isolated" and "purified" do not encompass nucleic acids or polypeptides present in their natural source.

The term "recombinant," when used herein to refer to a polypeptide or protein, means that a polypeptide or protein is derived from recombinant (e.g., microbial, insect, or mammalian) expression systems. "Microbial" refers to recombinant polypeptides or proteins made in bacterial or fungal (e.g., yeast) expression systems. As a product, "recombinant microbial" defines a polypeptide or protein essentially free of native endogenous substances and unaccompanied by associated native glycosylation. Polypeptides or proteins expressed in most bacterial cultures, e.g., E. coli, will be free of glycosylation modifications; polypeptides or proteins expressed in yeast will have a glycosylation pattern in general different from those expressed in mammalian cells.

The term "recombinant expression vehicle or vector" refers to a plasmid or phage or virus or vector, for expressing a polypeptide from a DNA (RNA) sequence. An expression vehicle can comprise a transcriptional unit comprising an assembly of (1) a genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers, (2) a structural or coding sequence which is transcribed into mRNA and translated into protein, and (3) appropriate transcription initiation and termination sequences. Structural units intended for use

in yeast or eukaryotic expression systems preferably include a leader sequence enabling extracellular secretion of translated protein by a host cell. Alternatively, where recombinant protein is expressed without a leader or transport sequence, it may include an amino terminal methionine residue. This residue may or may not be subsequently cleaved from the expressed recombinant protein to provide a final product.

The term "recombinant expression system" means host cells which have stably integrated a recombinant transcriptional unit into chromosomal DNA or carry the recombinant transcriptional unit extrachromosomally. Recombinant expression systems as defined herein will express heterologous polypeptides or proteins upon induction of the regulatory elements linked to the DNA segment or synthetic gene to be expressed. This term also means host cells which have stably integrated a recombinant genetic element or elements having a regulatory role in gene expression, for example, promoters or enhancers. Recombinant expression systems as defined herein will express polypeptides or proteins endogenous to the cell upon induction of the regulatory elements linked to the endogenous DNA segment or gene to be expressed. The cells can be prokaryotic or eukaryotic.

10

15

20

25

30

35

The term "secreted" includes a protein that is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence when it is expressed in a suitable host cell. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g., receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins that are transported across the membrane of the endoplasmic reticulum. "Secreted" proteins are also intended to include proteins containing non-typical signal sequences (e.g. Interleukin-1 Beta, see Krasney, P.A. and Young, P.R. (1992) Cytokine 4(2):134 -143) and factors released from damaged cells (e.g. Interleukin-1 Receptor Antagonist, see Arend, W.P. et. al. (1998) Annu. Rev. Immunol. 16:27-55)

Where desired, an expression vector may be designed to contain a "signal or leader sequence" which will direct the polypeptide through the membrane of a cell. Such a sequence may be naturally present on the polypeptides of the present invention or provided from heterologous protein sources by recombinant DNA techniques.

The term "stringent" is used to refer to conditions that are commonly understood in the art as stringent. Stringent conditions can include highly stringent conditions (i.e., hybridization to filter-bound DNA in 0.5 M NaHPO₄, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C, and washing in 0.1X SSC/0.1% SDS at 68°C), and moderately stringent conditions (i.e., washing in 0.2X SSC/0.1% SDS at 42°C). Other exemplary hybridization conditions are described herein in the examples.

In instances of hybridization of deoxyoligonucleotides, additional exemplary stringent hybridization conditions include washing in 6X SSC/0.05% sodium pyrophosphate at 37°C (for 14-base oligonucleotides), 48°C (for 17-base oligos), 55°C (for 20-base oligonucleotides), and 60°C (for 23-base oligonucleotides).

5

10

15

20

25

30

As used herein, "substantially equivalent" can refer both to nucleotide and amino acid sequences, for example a mutant sequence, that varies from a reference sequence by one or more substitutions, deletions, or additions, the net effect of which does not result in an adverse functional dissimilarity between the reference and subject sequences. Typically, such a substantially equivalent sequence varies from one of those listed herein by no more than about 35% (i.e., the number of individual residue substitutions, additions, and/or deletions in a substantially equivalent sequence, as compared to the corresponding reference sequence, divided by the total number of residues in the substantially equivalent sequence is about 0.35 or less). Such a sequence is said to have 65% sequence identity to the listed sequence. In one embodiment, a substantially equivalent, e.g., mutant, sequence of the invention varies from a listed sequence by no more than 30% (70% sequence identity); in a variation of this embodiment, by no more than 25% (75% sequence identity); and in a further variation of this embodiment, by no more than 20% (80% sequence identity) and in a further variation of this embodiment, by no more than 10% (90% sequence identity) and in a further variation of this embodiment, by no more that 5% (95% sequence identity). Substantially equivalent, e.g., mutant, amino acid sequences according to the invention preferably have at least 80% sequence identity with a listed amino acid sequence, more preferably at least 85% sequence identity, more preferably at least 90% sequence identity, more preferably at least 95% identity, more preferably at least 98% identity, and most preferably at least 99% identity. Substantially equivalent nucleotide sequences of the invention can have lower percent sequence identities, taking into account, for example, the redundancy or degeneracy of the genetic code. Preferably, nucleotide sequence has at least about 65% identity, more preferably at least about 75% identity, more preferably at least about 80% sequence identity, more preferably at least about 85% sequence identity, more preferably at least about 90% sequence identity, and most preferably at least about 95% identity, more preferably at least about 98% sequence identity, and most preferably at least about 99% sequence identity. For the purposes of the present invention, sequences having substantially equivalent biological activity and substantially equivalent expression characteristics are considered substantially equivalent. For the purposes of determining equivalence, truncation of the mature sequence (e.g., via a mutation which creates a spurious stop codon) should be disregarded. Sequence identity may be determined, e.g., using the Jotun Hein method (Hein, J.

(1990) Methods Enzymol. 183:626-645). Identity between sequences can also be determined by other methods known in the art, e.g. by varying hybridization conditions.

The term "totipotent" refers to the capability of a cell to differentiate into all of the cell types of an adult organism.

The term "transformation" means introducing DNA into a suitable host cell so that the DNA is replicable, either as an extrachromosomal element, or by chromosomal integration. The term "transfection" refers to the taking up of an expression vector by a suitable host cell, whether or not any coding sequences are in fact expressed. The term "infection" refers to the introduction of nucleic acids into a suitable host cell by use of a virus or viral vector.

As used herein, an "uptake modulating fragment," UMF, means a series of nucleotides which mediate the uptake of a linked DNA fragment into a cell. UMFs can be readily identified using known UMFs as a target sequence or target motif with the computer-based systems described below. The presence and activity of a UMF can be confirmed by attaching the suspected UMF to a marker sequence. The resulting nucleic acid molecule is then incubated with an appropriate host under appropriate conditions and the uptake of the marker sequence is determined. As described above, a UMF will increase the frequency of uptake of a linked marker sequence.

Each of the above terms is meant to encompass all that is described for each, unless the context dictates otherwise.

20

25

30

35

5

10

15

4.2 NUCLEIC ACIDS OF THE INVENTION

Nucleotide sequences of the invention are set forth in the Sequence Listing.

The isolated polynucleotides of the invention include a polynucleotide comprising the nucleotide sequences of SEQ ID NO:1-1350; a polynucleotide encoding any one of the peptide sequences of SEQ ID NO:1351-2700; and a polynucleotide comprising the nucleotide sequence encoding the mature protein coding sequence of the polypeptides of any one of SEQ ID NO:1351-2700. The polynucleotides of the present invention also include, but are not limited to, a polynucleotide that hybridizes under stringent conditions to (a) the complement of any of the nucleotides sequences of SEQ ID NO:1-1350; (b) nucleotide sequences encoding any one of the amino acid sequences set forth in the Sequence Listing; (c) a polynucleotide which is an allelic variant of any polynucleotide recited above; (d) a polynucleotide which encodes a species homolog of any of the proteins recited above; or (e) a polynucleotide that encodes a polypeptide comprising a specific domain or truncation of the polypeptides of SEQ ID NO: 1351-2700. Domains of interest may depend on the nature of the encoded polypeptide; e.g., domains in receptor-like polypeptides include ligand-binding, extracellular, transmembrane, or cytoplasmic

domains, or combinations thereof; domains in immunoglobulin-like proteins include the variable immunoglobulin-like domains; domains in enzyme-like polypeptides include catalytic and substrate binding domains; and domains in ligand polypeptides include receptor-binding domains.

The polynucleotides of the invention include naturally occurring or wholly or partially synthetic DNA, e.g., cDNA and genomic DNA, and RNA, e.g., mRNA. The polynucleotides may include all of the coding region of the cDNA or may represent a portion of the coding region of the cDNA.

5

10

15

20

25

30

The present invention also provides genes corresponding to the cDNA sequences disclosed herein. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. Further 5' and 3' sequence can be obtained using methods known in the art. For example, full length cDNA or genomic DNA that corresponds to any of the polynucleotides of SEQ ID NO:1-1350 can be obtained by screening appropriate cDNA or genomic DNA libraries under suitable hybridization conditions using any of the polynucleotides of SEQ ID NO:1-1350 or a portion thereof as a probe. Alternatively, the polynucleotides of SEQ ID NO:1-1350 may be used as the basis for suitable primer(s) that allow identification and/or amplification of genes in appropriate genomic DNA or cDNA libraries.

The nucleic acid sequences of the invention can be assembled from ESTs and sequences (including cDNA and genomic sequences) obtained from one or more public databases, such as dbEST, gbpri, and UniGene. The EST sequences can provide identifying sequence information, representative fragment or segment information, or novel segment information for the full-length gene.

The polynucleotides of the invention also provide polynucleotides including nucleotide sequences that are substantially equivalent to the polynucleotides recited above. Polynucleotides according to the invention can have, e.g., at least about 65%, at least about 70%, at least about 75%, at least about 80%, 81%, 82%, 83%, 84%, more typically at least about 85%, 86%, 87%, 88%, 89%, more typically at least about 90%, 91%, 92%, 93%, 94%, and even more typically at least about 95%, 96%, 97%, 98%, 99%, sequence identity to a polynucleotide recited above.

Included within the scope of the nucleic acid sequences of the invention are nucleic acid sequence fragments that hybridize under stringent conditions to any of the nucleotide sequences of SEQ ID NO:1-1350, or complements thereof, which fragment is greater than about 5 nucleotides, preferably 7 nucleotides, more preferably greater than 9 nucleotides and most preferably greater than 17 nucleotides. Fragments of, e.g. 15, 17, or 20 nucleotides or more that

are selective for (i.e. specifically hybridize to any one of the polynucleotides of the invention) are contemplated. Probes capable of specifically hybridizing to a polynucleotide can differentiate polynucleotide sequences of the invention from other polynucleotide sequences in the same family of genes or can differentiate human genes from genes of other species, and are preferably based on unique nucleotide sequences.

The sequences falling within the scope of the present invention are not limited to these specific sequences, but also include allelic and species variations thereof. Allelic and species variations can be routinely determined by comparing the sequence provided SEQ ID NO:1-1350, a representative fragment thereof, or a nucleotide sequence at least 90% identical, preferably 95% identical, to SEQ ID NO:1-1350 with a sequence from another isolate of the same species. Furthermore, to accommodate codon variability, the invention includes nucleic acid molecules coding for the same amino acid sequences as do the specific ORFs disclosed herein. In other words, in the coding region of an ORF, substitution of one codon for another codon that encodes the same amino acid is expressly contemplated.

10

15

20

25

30

The nearest neighbor or homology result for the nucleic acids of the present invention, including SEQ ID NO:1-1350, can be obtained by searching a database using an algorithm or a program. Preferably, a BLAST which stands for Basic Local Alignment Search Tool is used to search for local sequence alignments (Altshul, S.F. J Mol. Evol. 36 290-300 (1993) and Altschul S.F. et al. J. Mol. Biol. 21:403-410 (1990)). Alternatively a FASTA version 3 search against Genpept, using Fastxy algorithm.

Species homologs (or orthologs) of the disclosed polynucleotides and proteins are also provided by the present invention. Species homologs may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species.

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotide which also encode proteins which are identical, homologous or related to that encoded by the polynucleotides.

The nucleic acid sequences of the invention are further directed to sequences which encode variants of the described nucleic acids. These amino acid sequence variants may be prepared by methods known in the art by introducing appropriate nucleotide changes into a native or variant polynucleotide. There are two variables in the construction of amino acid sequence variants: the location of the mutation and the nature of the mutation. Nucleic acids encoding the amino acid sequence variants are preferably constructed by mutating the polynucleotide to encode an amino acid sequence that does not occur in nature. These nucleic

35

acid alterations can be made at sites that differ in the nucleic acids from different species (variable positions) or in highly conserved regions (constant regions). Sites at such locations will typically be modified in series, e.g., by substituting first with conservative choices (e.g., hydrophobic amino acid to a different hydrophobic amino acid) and then with more distant choices (e.g., hydrophobic amino acid to a charged amino acid), and then deletions or insertions may be made at the target site. Amino acid sequence deletions generally range from about 1 to 30 residues, preferably about 1 to 10 residues, and are typically contiguous. Amino acid insertions include amino- and/or carboxyl-terminal fusions ranging in length from one to one hundred or more residues, as well as intrasequence insertions of single or multiple amino acid residues. Intrasequence insertions may range generally from about 1 to 10 amino residues, preferably from 1 to 5 residues. Examples of terminal insertions include the heterologous signal sequences necessary for secretion or for intracellular targeting in different host cells and sequences such as FLAG or poly-histidine sequences useful for purifying the expressed protein.

In a preferred method, polynucleotides encoding the novel amino acid sequences are changed via site-directed mutagenesis. This method uses oligonucleotide sequences to alter a polynucleotide to encode the desired amino acid variant, as well as sufficient adjacent nucleotides on both sides of the changed amino acid to form a stable duplex on either side of the site of being changed. In general, the techniques of site-directed mutagenesis are well known to those of skill in the art and this technique is exemplified by publications such as, Edelman et al., DNA 2:183 (1983). A versatile and efficient method for producing site-specific changes in a polynucleotide sequence was published by Zoller and Smith, Nucleic Acids Res. 10:6487-6500 (1982). PCR may also be used to create amino acid sequence variants of the novel nucleic acids. When small amounts of template DNA are used as starting material, primer(s) that differs slightly in sequence from the corresponding region in the template DNA can generate the desired amino acid variant. PCR amplification results in a population of product DNA fragments that differ from the polynucleotide template encoding the polypeptide at the position specified by the primer. The product DNA fragments replace the corresponding region in the plasmid and this gives a polynucleotide encoding the desired amino acid variant.

15

20

25

30

A further technique for generating amino acid variants is the cassette mutagenesis technique described in Wells et al., Gene 34:315 (1985); and other mutagenesis techniques well known in the art, such as, for example, the techniques in Sambrook et al., supra, and Current Protocols in Molecular Biology, Ausubel et al. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be used in the practice of the invention for the cloning and expression

of these novel nucleic acids. Such DNA sequences include those which are capable of hybridizing to the appropriate novel nucleic acid sequence under stringent conditions.

5

10

15

20

25

30

35

Polynucleotides encoding preferred polypeptide truncations of the invention can be used to generate polynucleotides encoding chimeric or fusion proteins comprising one or more domains of the invention and heterologous protein sequences.

The polynucleotides of the invention additionally include the complement of any of the polynucleotides recited above. The polynucleotide can be DNA (genomic, cDNA, amplified, or synthetic) or RNA. Methods and algorithms for obtaining such polynucleotides are well known to those of skill in the art and can include, for example, methods for determining hybridization conditions that can routinely isolate polynucleotides of the desired sequence identities.

In accordance with the invention, polynucleotide sequences comprising the mature protein coding sequences corresponding to any one of SEQ ID NO:1-1350, or functional equivalents thereof, may be used to generate recombinant DNA molecules that direct the expression of that nucleic acid, or a functional equivalent thereof, in appropriate host cells. Also included are the cDNA inserts of any of the clones identified herein.

A polynucleotide according to the invention can be joined to any of a variety of other nucleotide sequences by well-established recombinant DNA techniques (see Sambrook J et al. (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, NY). Useful nucleotide sequences for joining to polynucleotides include an assortment of vectors, e.g., plasmids, cosmids, lambda phage derivatives, phagemids, and the like, that are well known in the art. Accordingly, the invention also provides a vector including a polynucleotide of the invention and a host cell containing the polynucleotide. In general, the vector contains an origin of replication functional in at least one organism, convenient restriction endonuclease sites, and a selectable marker for the host cell. Vectors according to the invention include expression vectors, replication vectors, probe generation vectors, and sequencing vectors. A host cell according to the invention can be a prokaryotic or eukaryotic cell and can be a unicellular organism or part of a multicellular organism.

The present invention further provides recombinant constructs comprising a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof or any other polynucleotides of the invention. In one embodiment, the recombinant constructs of the present invention comprise a vector, such as a plasmid or viral vector, into which a nucleic acid having any of the nucleotide sequences of SEQ ID NO:1-1350 or a fragment thereof is inserted, in a forward or reverse orientation. In the case of a vector comprising one of the ORFs of the present invention, the vector may further comprise regulatory sequences, including for example, a promoter, operably linked to the ORF. Large numbers of suitable vectors and promoters are

known to those of skill in the art and are commercially available for generating the recombinant constructs of the present invention. The following vectors are provided by way of example. Bacterial: pBs, phagescript, PsiX174, pBluescript SK, pBs KS, pNH8a, pNH16a, pNH18a, pNH46a (Stratagene); pTrc99A, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia). Eukaryotic: pWLneo, pSV2cat, pOG44, PXTI, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia).

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman et al., *Nucleic Acids Res.* 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, *Methods in Enzymology* 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

10

15

20

25

30

35

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacl, lacZ, T3, T7, gpt, lambda PR, and trc. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), a-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an amino terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product. Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or

more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within the host. Suitable prokaryotic hosts for transformation include *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium* and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

As a representative but non-limiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017). Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM 1 (Promega Biotech, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed. Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced or derepressed by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period. Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Polynucleotides of the invention can also be used to induce immune responses. For example, as described in Fan et al., *Nat. Biotech.* 17:870-872 (1999), incorporated herein by reference, nucleic acid sequences encoding a polypeptide may be used to generate antibodies against the encoded polypeptide following topical administration of naked plasmid DNA or following injection, and preferably intramuscular injection of the DNA. The nucleic acid sequences are preferably inserted in a recombinant expression vector and may be in the form of naked DNA.

25 4.3 ANTISENSE

10

15

20

30

Another aspect of the invention pertains to isolated antisense nucleic acid molecules that are hybridizable to or complementary to the nucleic acid molecule comprising the nucleotide sequence of SEQ ID NO:1-1350, or fragments, analogs or derivatives thereof. An "antisense" nucleic acid comprises a nucleotide sequence that is complementary to a "sense" nucleic acid encoding a protein, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. In specific aspects, antisense nucleic acid molecules are provided that comprise a sequence complementary to at least about 10, 25, 50, 100, 250 or 500 nucleotides or an entire coding strand, or to only a portion thereof. Nucleic acid molecules encoding fragments, homologs, derivatives and analogs of a protein of any of SEQ ID

NO:1351-2700 or antisense nucleic acids complementary to a nucleic acid sequence of SEQ ID NO:1-1350 are additionally provided.

In one embodiment, an antisense nucleic acid molecule is antisense to a "coding region" of the coding strand of a nucleotide sequence of the invention. The term "coding region" refers to the region of the nucleotide sequence comprising codons which are translated into amino acid residues. In another embodiment, the antisense nucleic acid molecule is antisense to a "noncoding region" of the coding strand of a nucleotide sequence of the invention. The term "noncoding region" refers to 5' and 3' sequences which flank the coding region that are not translated into amino acids (i.e., also referred to as 5' and 3' untranslated regions).

10

15

20

25

30

35

Given the coding strand sequences encoding a nucleic acid disclosed herein (e.g., SEQ ID NO:1-1350), antisense nucleic acids of the invention can be designed according to the rules of Watson and Crick or Hoogsteen base pairing. The antisense nucleic acid molecule can be complementary to the entire coding region of a mRNA, but more preferably is an oligonucleotide that is antisense to only a portion of the coding or noncoding region of a mRNA. For example, the antisense oligonucleotide can be complementary to the region surrounding the translation start site of a mRNA. An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis or enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used.

Examples of modified nucleotides that can be used to generate the antisense nucleic acid include: 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the

antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

5

15

20

30

35

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated in situ such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a protein according to the invention to thereby inhibit expression of the protein, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule that binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies that bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of antisense molecules, vector constructs in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

In yet another embodiment, the antisense nucleic acid molecule of the invention is an -a nomeric nucleic acid molecule. An -a nomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual -units, the strands run parallel to each other (Gaultier et al. (1987) Nucleic Acids Res 15: 6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) Nucleic Acids Res 15: 6131-6148) or a chimeric RNA -DNA analogue (Inoue et al. (1987) FEBS Lett 215: 327-330).

4.4 RIBOZYMES AND PNA MOIETIES

In still another embodiment, an antisense nucleic acid of the invention is a ribozyme. Ribozymes are catalytic RNA molecules with ribonuclease activity that are capable of cleaving a single-stranded nucleic acid, such as a mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) Nature 334:585-591)) can be used to catalytically cleave a mRNA transcripts to thereby inhibit translation of a mRNA. A ribozyme having specificity for a nucleic acid of the invention can be

designed based upon the nucleotide sequence of a DNA disclosed herein (i.e., SEQ ID NO:1-1350). For example, a derivative of a Tetrahymena L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a SECX-encoding mRNA. See, e.g., Cech et al. U.S. Pat. No. 4,987,071; and Cech et al. U.S. Pat. No. 5,116,742. Alternatively, SECX mRNA can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. See, e.g., Bartel et al., (1993) Science 261:1411-1418.

Alternatively, gene expression can be inhibited by targeting nucleotide sequences complementary to the regulatory region (e.g., promoter and/or enhancers) to form triple helical structures that prevent transcription of the gene in target cells. See generally, Helene. (1991)

Anticancer Drug Des. 6: 569-84; Helene. et al. (1992) Ann. N.Y. Acad. Sci. 660:27-36; and Maher (1992) Bioassays 14: 807-15.

In various embodiments, the nucleic acids of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (see Hyrup et al. (1996) Bioorg Med Chem 4: 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al. (1996) above; Perry-O'Keefe et al. (1996) PNAS 93: 14670-675.

. 15

20

25

30

35

PNAs of the invention can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs of the invention can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (Hyrup B. (1996) above); or as probes or primers for DNA sequence and hybridization (Hyrup et al. (1996), above; Perry-O'Keefe (1996), above).

In another embodiment, PNAs of the invention can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated that may

combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNase H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996) above). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996) above and Finn et al. (1996) Nucl Acids Res 24: 3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry, and modified nucleoside analogs, e.g., 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite, can be used between the PNA and the 5' end of DNA (Mag et al. (1989) Nucl Acid Res 17: 5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) above). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment. See, Petersen et al. (1975) Bioorg Med Chem Lett 5: 1119-11124.

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., 1989, Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556;

Lemaitre et al., 1987, Proc. Natl. Acad. Sci. 84:648-652; PCT Publication No. W088/09810) or the blood-brain barrier (see, e.g., PCT Publication No. W089/10134). In addition,

20 oligonucleotides can be modified with hybridization triggered cleavage agents (See, e.g., Krol et al., 1988, BioTechniques 6:958-976) or intercalating agents. (See, e.g., Zon, 1988, Pharm. Res. 5: 539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, a hybridization triggered cross-linking agent, a transport agent, a hybridization-triggered cleavage agent, etc.

25

30

35

10

4.5 HOSTS

The present invention further provides host cells genetically engineered to contain the polynucleotides of the invention. For example, such host cells may contain nucleic acids of the invention introduced into the host cell using known transformation, transfection or infection methods. The present invention still further provides host cells genetically engineered to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in the cell.

Knowledge of nucleic acid sequences allows for modification of cells to permit, or increase, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous

recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the polypeptide at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the encoding sequences. See, for example, PCT International Publication No. WO94/12650, PCT International Publication No. WO92/20808, and PCT International Publication No. WO91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the coding sequence, amplification of the marker DNA by standard selection methods results in coamplification of the desired protein coding sequences in the cells.

10

15

20

25

30

35

The host cell can be a higher eukaryotic host cell, such as a mammalian cell, a lower eukaryotic host cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the recombinant construct into the host cell can be effected by calcium phosphate transfection, DEAE, dextran mediated transfection, or electroporation (Davis, L. et al., *Basic Methods in Molecular Biology* (1986)). The host cells containing one of the polynucleotides of the invention, can be used in conventional manners to produce the gene product encoded by the isolated fragment (in the case of an ORF) or can be used to produce a heterologous protein under the control of the EMF.

Any host/vector system can be used to express one or more of the ORFs of the present invention. These include, but are not limited to, eukaryotic hosts such as HeLa cells, Cv-1 cell, COS cells, 293 cells, and Sf9 cells, as well as prokaryotic host such as *E. coli* and *B. subtilis*. The most preferred cells are those which do not normally express the particular polypeptide or protein or which expresses the polypeptide or protein at low natural level. Mature proteins can be expressed in mammalian cells, yeast, bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., in Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, New York (1989), the disclosure of which is hereby incorporated by reference.

Various mammalian cell culture systems can also be employed to express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, Cell 23:175 (1981). Other cell lines capable of expressing a compatible vector are, for example, the C127, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3

cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from *in vitro* culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 viral genome, for example, SV40 origin, early promoter, enhancer, splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements. Recombinant polypeptides and proteins produced in bacterial culture are usually isolated by initial extraction from cell pellets, followed by one or more salting-out, aqueous ion exchange or size exclusion chromatography steps. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps. Microbial cells employed in expression of proteins can be disrupted by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or insects or in prokaryotes such as bacteria. Potentially suitable yeast strains include Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces strains, Candida, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include Escherichia coli, Bacillus subtilis, Salmonella typhimurium, or any bacterial strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequence include polyadenylation signals, mRNA stability elements, splice

sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which alter or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA. allowing for the selection of cells in which the exogenous DNA has integrated into the host cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.6 POLYPEPTIDES OF THE INVENTION

10

15

20

25

30

35

The isolated polypeptides of the invention include, but are not limited to, a polypeptide comprising: the amino acid sequences set forth as any one of SEQ ID NO:1351-2700 or an amino acid sequence encoded by any one of the nucleotide sequences SEQ ID NO:1-1350 or the corresponding full length or mature protein. Polypeptides of the invention also include polypeptides preferably with biological or immunological activity that are encoded by: (a) a polynucleotide having any one of the nucleotide sequences set forth in SEO ID NO:1-1350 or (b)

polynucleotides encoding any one of the amino acid sequences set forth as SEQ ID NO:1351-2700 or (c) polynucleotides that hybridize to the complement of the polynucleotides of either (a) or (b) under stringent hybridization conditions. The invention also provides biologically active or immunologically active variants of any of the amino acid sequences set forth as SEQ ID NO:1351-2700 or the corresponding full length or mature protein; and "substantial equivalents" thereof (e.g., with at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, 86%, 87%, 88%, 89%, at least about 90%, 91%, 92%, 93%, 94%, typically at least about 95%, 96%, 97%, more typically at least about 98%, or most typically at least about 99% amino acid identity) that retain biological activity. Polypeptides encoded by allelic variants may have a similar, increased, or decreased activity compared to polypeptides comprising SEQ ID NO:1351-2700.

10

15

20

25

30

35

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H. U. Saragovi, et al., Bio/Technology 10, 773-778 (1992) and in R. S. McDowell, et al., J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites.

The present invention also provides both full-length and mature forms (for example, without a signal sequence or precursor sequence) of the disclosed proteins. The protein coding sequence is identified in the sequence listing by translation of the disclosed nucleotide sequences. The mature form of such protein may be obtained by expression of a full-length polynucleotide in a suitable mammalian cell or other host cell. The sequence of the mature form of the protein is also determinable from the amino acid sequence of the full-length form. Where proteins of the present invention are membrane bound, soluble forms of the proteins are also provided. In such forms, part or all of the regions causing the proteins to be membrane bound are deleted so that the proteins are fully secreted from the cell in which they are expressed.

Protein compositions of the present invention may further comprise an acceptable carrier, such as a hydrophilic, e.g., pharmaceutically acceptable, carrier.

The present invention further provides isolated polypeptides encoded by the nucleic acid fragments of the present invention or by degenerate variants of the nucleic acid fragments of the present invention. By "degenerate variant" is intended nucleotide fragments which differ from a nucleic acid fragment of the present invention (e.g., an ORF) by nucleotide sequence but, due to the degeneracy of the genetic code, encode an identical polypeptide sequence. Preferred nucleic acid fragments of the present invention are the ORFs that encode proteins.

A variety of methodologies known in the art can be utilized to obtain any one of the isolated polypeptides or proteins of the present invention. At the simplest level, the amino acid sequence can be synthesized using commercially available peptide synthesizers. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. This technique is particularly useful in producing small peptides and fragments of larger polypeptides. Fragments are useful, for example, in generating antibodies against the native polypeptide. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

10

15

20

30

35

The polypeptides and proteins of the present invention can alternatively be purified from cells which have been altered to express the desired polypeptide or protein. As used herein, a cell is said to be altered to express a desired polypeptide or protein when the cell, through genetic manipulation, is made to produce a polypeptide or protein which it normally does not produce or which the cell normally produces at a lower level. One skilled in the art can readily adapt procedures for introducing and expressing either recombinant or synthetic sequences into eukaryotic or prokaryotic cells in order to generate a cell which produces one of the polypeptides or proteins of the present invention.

The invention also relates to methods for producing a polypeptide comprising growing a culture of host cells of the invention in a suitable culture medium, and purifying the protein from the cells or the culture in which the cells are grown. For example, the methods of the invention include a process for producing a polypeptide in which a host cell containing a suitable expression vector that includes a polynucleotide of the invention is cultured under conditions that allow expression of the encoded polypeptide. The polypeptide can be recovered from the culture, conveniently from the culture medium, or from a lysate prepared from the host cells and further purified. Preferred embodiments include those in which the protein produced by such process is a full length or mature form of the protein.

In an alternative method, the polypeptide or protein is purified from bacterial cells which naturally produce the polypeptide or protein. One skilled in the art can readily follow known methods for isolating polypeptides and proteins in order to obtain one of the isolated polypeptides or proteins of the present invention. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography, and immuno-affinity chromatography. See, e.g., Scopes, Protein Purification: Principles and Practice, Springer-Verlag (1994); Sambrook, et al., in Molecular Cloning: A Laboratory Manual; Ausubel et al., Current Protocols in Molecular Biology. Polypeptide fragments that

retain biological/immunological activity include fragments comprising greater than about 100 amino acids, or greater than about 200 amino acids, and fragments that encode specific protein domains.

5

10

15

20

25

30

35

The purified polypeptides can be used in *in vitro* binding assays which are well known in the art to identify molecules which bind to the polypeptides. These molecules include but are not limited to, for *e.g.*, small molecules, molecules from combinatorial libraries, antibodics or other proteins. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

In addition, the peptides of the invention or molecules capable of binding to the peptides may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for SEQ ID NO:1351-2700.

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications, in the peptide or DNA sequence, can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Pat. No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein. Regions of the protein that are important for the protein function can be determined by various methods known in the art including the alanine-scanning method which involved systematic substitution of single or strings of amino acids with alanine, followed by testing the resulting alanine-containing variant for biological activity. This type of analysis determines the importance of the substituted amino acid(s) in biological activity. Regions of the protein that are important for protein function may be determined by the eMATRIX program.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and are useful for screening or other immunological

methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are encompassed by the present invention.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, e.g., Invitrogen, San Diego, Calif., U.S.A. (the MaxBatTM kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

5

10

15

20

25

30

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearlTM or Cibacrom blue 3GA SepharoseTM; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX), or as a His tag. Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, Mass.), Pharmacia (Piscataway, N.J.) and Invitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("FLAG®") is commercially available from Kodak (New Haven, Conn.).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The polypeptides of the invention include analogs (variants). This embraces fragments, as well as peptides in which one or more amino acids has been deleted, inserted, or substituted. Also, analogs of the polypeptides of the invention embrace fusions of the polypeptides or modifications of the polypeptides of the invention, wherein the polypeptide or analog is fused to another moiety or moieties, *e.g.*, targeting moiety or another therapeutic agent. Such analogs may exhibit improved properties such as activity and/or stability. Examples of moieties which may be fused to the polypeptide or an analog include, for example, targeting moieties which provide for the delivery of polypeptide to pancreatic cells, *e.g.*, antibodies to pancreatic cells, antibodies to immune cells such as T-cells, monocytes, dendritic cells, granulocytes, etc., as well as receptor and ligands expressed on pancreatic or immune cells. Other moieties which may be fused to the polypeptide include therapeutic agents which are used for treatment, for example, immunosuppressive drugs such as cyclosporin, SK506, azathioprine, CD3 antibodies and steroids. Also, polypeptides may be fused to immune modulators, and other cytokines such as alpha or beta interferon.

15

20

25

30

35

10

4.6.1 DETERMINING POLYPEPTIDE AND POLYNUCLEOTIDE IDENTITY AND SIMILARITY

Preferred identity and/or similarity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in computer programs including, but are not limited to, the GCG program package, including GAP (Devereux, J., et al., Nucleic Acids Research 12(1):387 (1984); Genetics Computer Group, University of Wisconsin, Madison, WI), BLASTP, BLASTN, BLASTX, FASTA (Altschul, S.F. et al., J. Molec. Biol. 215:403-410 (1990), PSI-BLAST (Altschul S.F. et al., Nucleic Acids Res. vol. 25, pp. 3389-3402, herein incorporated by reference), eMatrix software (Wu et al., J. Comp. Biol., Vol. 6, pp. 219-235 (1999), herein incorporated by reference), eMotif software (Nevill-Manning et al, ISMB-97, Vol. 4, pp. 202-209, herein incorporated by reference), pFam software (Sonnhammer et al., Nucleic Acids Res., Vol. 26(1), pp. 320-322 (1998), herein incorporated by reference) and the Kyte-Doolittle hydrophobocity prediction algorithm (J. Mol Biol, 157, pp. 105-31 (1982), incorporated herein by reference). The BLAST programs are publicly available from the National Center for Biotechnology Information (NCBI) and other sources (BLAST Manual, Altschul, S., et al., NCB NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215:403-410 (1990).

4.7 CHIMERIC AND FUSION PROTEINS

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises a polypeptide of the invention operatively linked to

another polypeptide. Within a fusion protein the polypeptide according to the invention can correspond to all or a portion of a protein according to the invention. In one embodiment, a fusion protein comprises at least one biologically active portion of a protein according to the invention. In another embodiment, a fusion protein comprises at least two biologically active portions of a protein according to the invention. Within the fusion protein, the term "operatively linked" is intended to indicate that the polypeptide according to the invention and the other polypeptide are fused in-frame to each other. The polypeptide can be fused to the N-terminus or C-terminus.

For example, in one embodiment a fusion protein comprises a polypeptide according to the invention operably linked to the extracellular domain of a second protein.

10

15

20

25

30

35

In another embodiment, the fusion protein is a GST-fusion protein in which the polypeptide sequences of the invention are fused to the C-terminus of the GST (i.e., glutathione S-transferase) sequences.

In another embodiment, the fusion protein is an immunoglobulin fusion protein in which the polypeptide sequences according to the invention comprises one or more domains are fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand and a protein of the invention on the surface of a cell, to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion proteins can be used to affect the bioavailability of a cognate ligand. Inhibition of the ligand/protein interaction may be useful therapeutically for both the treatment of proliferative and differentiative disorders, *e.g.*, cancer as well as modulating (*e.g.*, promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies in a subject, to purify ligands, and in screening assays to identify molecules that inhibit the interaction of a polypeptide of the invention with a ligand.

A chimeric or fusion protein of the invention can be produced by standard recombinant DNA techniques. For example, DNA fragments coding for the different polypeptide sequences are ligated together in-frame in accordance with conventional techniques, e.g., by employing blunt-ended or stagger-ended termini for ligation, restriction enzyme digestion to provide for appropriate termini, filling-in of cohesive ends as appropriate, alkaline phosphatase treatment to avoid undesirable joining, and enzymatic ligation. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers.

Alternatively, PCR amplification of gene fragments can be carried out using anchor primers that give rise to complementary overhangs between two consecutive gene fragments that can subsequently be annealed and reamplified to generate a chimeric gene sequence (see, for

example, Ausubel et al. (eds.) CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, John Wiley & Sons, 1992). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the protein of the invention.

4.8 GENE THERAPY

5

10

15

20

25

30

Mutations in the polynucleotides of the invention gene may result in loss of normal function of the encoded protein. The invention thus provides gene therapy to restore normal activity of the polypeptides of the invention; or to treat disease states involving polypeptides of the invention. Delivery of a functional gene encoding polypeptides of the invention to appropriate cells is effected ex vivo, in situ, or in vivo by use of vectors, and more particularly viral vectors (e.g., adenovirus, adeno-associated virus, or a retrovirus), or ex vivo by use of physical DNA transfer methods (e.g., liposomes or chemical treatments). See, for example, Anderson, Nature, supplement to vol. 392, no. 6679, pp.25-20 (1998). For additional reviews of gene therapy technology see Friedmann, Science, 244: 1275-1281 (1989); Verma, Scientific American: 68-84 (1990); and Miller, Nature, 357: 455-460 (1992). Introduction of any one of the nucleotides of the present invention or a gene encoding the polypeptides of the present invention can also be accomplished with extrachromosomal substrates (transient expression) or artificial chromosomes (stable expression). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes. Alternatively, it is contemplated that in other human disease states, preventing the expression of or inhibiting the activity of polypeptides of the invention will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of polypeptides of the invention.

Other methods inhibiting expression of a protein include the introduction of antisense molecules to the nucleic acids of the present invention, their complements, or their translated RNA sequences, by methods known in the art. Further, the polypeptides of the present invention can be inhibited by using targeted deletion methods, or the insertion of a negative regulatory element such as a silencer, which is tissue specific.

The present invention still further provides cells genetically engineered in vivo to express the polynucleotides of the invention, wherein such polynucleotides are in operative association with a regulatory sequence heterologous to the host cell which drives expression of the polynucleotides in

the cell. These methods can be used to increase or decrease the expression of the polynucleotides of the present invention.

Knowledge of DNA sequences provided by the invention allows for modification of cells to permit, increase, or decrease, expression of endogenous polypeptide. Cells can be modified (e.g., by homologous recombination) to provide increased polypeptide expression by replacing, in whole or in part, the naturally occurring promoter with all or part of a heterologous promoter so that the cells express the protein at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to the desired protein encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. It is also contemplated that, in addition to heterologous promoter DNA, amplifiable marker DNA (e.g., ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the desired protein coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the desired protein coding sequences in the cells.

10

15

20

25

30

35

In another embodiment of the present invention, cells and tissues may be engineered to express an endogenous gene comprising the polynucleotides of the invention under the control of inducible regulatory elements, in which case the regulatory sequences of the endogenous gene may be replaced by homologous recombination. As described herein, gene targeting can be used to replace a gene's existing regulatory region with a regulatory sequence isolated from a different gene or a novel regulatory sequence synthesized by genetic engineering methods. Such regulatory sequences may be comprised of promoters, enhancers, scaffold-attachment regions, negative regulatory elements, transcriptional initiation sites, regulatory protein binding sites or combinations of said sequences. Alternatively, sequences which affect the structure or stability of the RNA or protein produced may be replaced, removed, added, or otherwise modified by targeting. These sequences include polyadenylation signals, mRNA stability elements, splice sites, leader sequences for enhancing or modifying transport or secretion properties of the protein, or other sequences which after or improve the function or stability of protein or RNA molecules.

The targeting event may be a simple insertion of the regulatory sequence, placing the gene under the control of the new regulatory sequence, e.g., inserting a new promoter or enhancer or both upstream of a gene. Alternatively, the targeting event may be a simple deletion of a regulatory element, such as the deletion of a tissue-specific negative regulatory element. Alternatively, the targeting event may replace an existing element; for example, a tissue-specific enhancer can be replaced by an enhancer that has broader or different cell-type specificity than the naturally occurring elements. Here, the naturally occurring sequences are deleted and new sequences are

added. In all cases, the identification of the targeting event may be facilitated by the use of one or more selectable marker genes that are contiguous with the targeting DNA, allowing for the selection of cells in which the exogenous DNA has integrated into the cell genome. The identification of the targeting event may also be facilitated by the use of one or more marker genes exhibiting the property of negative selection, such that the negatively selectable marker is linked to the exogenous DNA, but configured such that the negatively selectable marker flanks the targeting sequence, and such that a correct homologous recombination event with sequences in the host cell genome does not result in the stable integration of the negatively selectable marker. Markers useful for this purpose include the Herpes Simplex Virus thymidine kinase (TK) gene or the bacterial xanthine-guanine phosphoribosyl-transferase (gpt) gene.

The gene targeting or gene activation techniques which can be used in accordance with this aspect of the invention are more particularly described in U.S. Patent No. 5,272,071 to Chappel; U.S. Patent No. 5,578,461 to Sherwin et al.; International Application No. PCT/US92/09627 (WO93/09222) by Selden et al.; and International Application No. PCT/US90/06436 (WO91/06667) by Skoultchi et al., each of which is incorporated by reference herein in its entirety.

4.9 TRANSGENIC ANIMALS

10

15

20

25

30

35

In preferred methods to determine biological functions of the polypeptides of the invention in vivo, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of a promoter of the polynucleotides of the invention is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous

promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

The polynucleotides of the present invention also make possible the development, through, e.g., homologous recombination or knock out strategies, of animals that fail to express polypeptides of the invention or that express a variant polypeptide. Such animals are useful as models for studying the *in vivo* activities of polypeptide as well as for studying modulators of the polypeptides of the invention.

In preferred methods to determine biological functions of the polypeptides of the invention *in vivo*, one or more genes provided by the invention are either over expressed or inactivated in the germ line of animals using homologous recombination [Capecchi, Science 244:1288-1292 (1989)]. Animals in which the gene is over expressed, under the regulatory control of exogenous or endogenous promoter elements, are known as transgenic animals. Animals in which an endogenous gene has been inactivated by homologous recombination are referred to as "knockout" animals. Knockout animals, preferably non-human mammals, can be prepared as described in U.S. Patent No. 5,557,032, incorporated herein by reference. Transgenic animals are useful to determine the roles polypeptides of the invention play in biological processes, and preferably in disease states. Transgenic animals are useful as model systems to identify compounds that modulate lipid metabolism. Transgenic animals, preferably non-human mammals, are produced using methods as described in U.S. Patent No 5,489,743 and PCT Publication No. WO94/28122, incorporated herein by reference.

Transgenic animals can be prepared wherein all or part of the polynucleotides of the invention promoter is either activated or inactivated to alter the level of expression of the polypeptides of the invention. Inactivation can be carried out using homologous recombination methods described above. Activation can be achieved by supplementing or even replacing the homologous promoter to provide for increased protein expression. The homologous promoter can be supplemented by insertion of one or more heterologous enhancer elements known to confer promoter activation in a particular tissue.

4.10 USES AND BIOLOGICAL ACTIVITY

10

15

20

30

35

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified herein. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA). The mechanism underlying the particular condition or pathology will dictate whether the

polypeptides of the invention, the polynucleotides of the invention or modulators (activators or inhibitors) thereof would be beneficial to the subject in need of treatment. Thus, "therapeutic compositions of the invention" include compositions comprising isolated polynucleotides (including recombinant DNA molecules, cloned genes and degenerate variants thereof) or polypeptides of the invention (including full length protein, mature protein and truncations or domains thereof), or compounds and other substances that modulate the overall activity of the target gene products, either at the level of target gene/protein expression or target protein activity. Such modulators include polypeptides, analogs, (variants), including fragments and fusion proteins, antibodies and other binding proteins; chemical compounds that directly or indirectly activate or inhibit the polypeptides of the invention (identified, e.g., via drug screening assays as described herein); antisense polynucleotides and polynucleotides suitable for triple helix formation; and in particular antibodies or other binding partners that specifically recognize one or more epitopes of the polypeptides of the invention.

The polypeptides of the present invention may likewise be involved in cellular activation or in one of the other physiological pathways described herein.

4.10.1 RESEARCH USES AND UTILITIES

10

15

20

25

30

35

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, that described in Gyuris et al., Cell 75:791-803 (1993)) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The polypeptides provided by the present invention can similarly be used in assays to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding polypeptide is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E. F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S. L. and A. R. Kimmel eds., 1987.

4.10.2 NUTRITIONAL USES

5

10

15

20

25

30

35

Polynucleotides and polypeptides of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the polypeptide or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the polypeptide or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

4.10.3 CYTOKINE AND CELL PROLIFERATION/DIFFERENTIATION ACTIVITY

A polypeptide of the present invention may exhibit activity relating to cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor-dependent cell proliferation assays, and hence the assays serve as a convenient

confirmation of cytokine activity. The activity of therapeutic compositions of the present invention is evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+(preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e, CMK, HUVEC, and Caco. Therapeutic compositions of the invention can be used in the following:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., I. Immunol. 149:3778-3783, 1992; Bowman et al., I. Immunol. 152:1756-1761, 1994.

10

15

20

30

35

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A. M. and Shevach, E. M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human interleukin-y, Schreiber, R. D. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L. S. and Lipsky, P. E. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Mcd. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse 25 and human interleukin 6--Nordan, R. In Current Protocols in Immunology, J. E. Coligan eds, Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Aced. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11--Bennett, F., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9-Ciarletta, A., Giannotti, J., Clark, S. C. and Turner, K. J. In Current Protocols in Immunology. J. E. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W Strober,

PCT/US01/03800 WO 01/57188

Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol, 140:508-512, 1988.

4.10.4 STEM CELL GROWTH FACTOR ACTIVITY

5

10

20

25

35

A polypeptide of the present invention may exhibit stem cell growth factor activity and be involved in the proliferation, differentiation and survival of pluripotent and totipotent stem cells including primordial germ cells, embryonic stem cells, hematopoietic stem cells and/or germ line stem cells. Administration of the polypeptide of the invention to stem cells in vivo or ex vivo is expected to maintain and expand cell populations in a totipotential or pluripotential state which would be useful for re-engineering damaged or diseased tissues, transplantation, manufacture of bio-pharmaceuticals and the development of bio-sensors. The ability to produce 15 large quantities of human cells has important working applications for the production of human proteins which currently must be obtained from non-human sources or donors, implantation of cells to treat diseases such as Parkinson's, Alzheimer's and other neurodegenerative diseases; tissues for grafting such as bone marrow, skin, cartilage, tendons, bone, muscle (including cardiac muscle), blood vessels, cornea, neural cells, gastrointestinal cells and others; and organs for transplantation such as kidney, liver, pancreas (including islet cells), heart and lung.

It is contemplated that multiple different exogenous growth factors and/or cytokines may be administered in combination with the polypeptide of the invention to achieve the desired effect, including any of the growth factors listed herein, other stem cell maintenance factors, and specifically including stem cell factor (SCF), leukemia inhibitory factor (LIF), Flt-3 ligand (Flt-3L), any of the interleukins, recombinant soluble IL-6 receptor fused to IL-6, macrophage inflammatory protein 1-alpha (MIP-1-alpha), G-CSF, GM-CSF, thrombopoietin (TPO), platelet factor 4 (PF-4), platelet-derived growth factor (PDGF), neural growth factors and basic fibroblast growth factor (bFGF).

Since totipotent stem cells can give rise to virtually any mature cell type, expansion of these cells in culture will facilitate the production of large quantities of mature cells. Techniques for culturing stem cells are known in the art and administration of polypeptides of the invention. optionally with other growth factors and/or cytokines, is expected to enhance the survival and proliferation of the stem cell populations. This can be accomplished by direct administration of the polypeptide of the invention to the culture medium. Alternatively, stroma cells transfected with a polynucleotide that encodes for the polypeptide of the invention can be used as a feeder

layer for the stem cell populations in culture or in vivo. Stromal support cells for feeder layers may include embryonic bone marrow fibroblasts, bone marrow stromal cells, fetal liver cells, or cultured embryonic fibroblasts (see U.S. Patent No. 5,690,926).

Stem cells themselves can be transfected with a polynucleotide of the invention to induce autocrine expression of the polypeptide of the invention. This will allow for generation of undifferentiated totipotential/pluripotential stem cell lines that are useful as is or that can then be differentiated into the desired mature cell types. These stable cell lines can also serve as a source of undifferentiated totipotential/pluripotential mRNA to create cDNA libraries and templates for polymerase chain reaction experiments. These studies would allow for the isolation and identification of differentially expressed genes in stem cell populations that regulate stem cell proliferation and/or maintenance.

10

15

20

25

30

35

Expansion and maintenance of totipotent stem cell populations will be useful in the treatment of many pathological conditions. For example, polypeptides of the present invention may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease, accidental damage or genetic disorders. The polypeptide of the invention may be useful for inducing the proliferation of neural cells and for the regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders which involve degeneration, death or trauma to neural cells or nerve tissue. In addition, the expanded stem cell populations can also be genetically altered for gene therapy purposes and to decrease host rejection of replacement tissues after grafting or implantation.

Expression of the polypeptide of the invention and its effect on stem cells can also be manipulated to achieve controlled differentiation of the stem cells into more differentiated cell types. A broadly applicable method of obtaining pure populations of a specific differentiated cell type from undifferentiated stem cell populations involves the use of a cell-type specific promoter driving a selectable marker. The selectable marker allows only cells of the desired type to survive. For example, stem cells can be induced to differentiate into cardiomyocytes (Wobus et al., Differentiation, 48: 173-182, (1991); Klug et al., J. Clin. Invest., 98(1): 216-224, (1998)) or skeletal muscle cells (Browder, L. W. In: *Principles of Tissue Engineering eds.* Lanza et al., Academic Press (1997)). Alternatively, directed differentiation of stem cells can be accomplished by culturing the stem cells in the presence of a differentiation factor such as retinoic acid and an antagonist of the polypeptide of the invention which would inhibit the effects of endogenous stem cell factor activity and allow differentiation to proceed.

In vitro cultures of stem cells can be used to determine if the polypeptide of the invention exhibits stem cell growth factor activity. Stem cells are isolated from any one of various cell

sources (including hematopoietic stem cells and embryonic stem cells) and cultured on a feeder layer, as described by Thompson et al. Proc. Natl. Acad. Sci, U.S.A., 92: 7844-7848 (1995), in the presence of the polypeptide of the invention alone or in combination with other growth factors or cytokines. The ability of the polypeptide of the invention to induce stem cells proliferation is determined by colony formation on semi-solid support *e.g.* as described by Bernstein et al., Blood, 77: 2316-2321 (1991).

4.10.5 HEMATOPOIESIS REGULATING ACTIVITY

5

30

35

A polypeptide of the present invention may be involved in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell disorders. Even marginal 10 biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid 15 cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or 20 treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and 25 paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either in-vivo or ex-vivo (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

Therapeutic compositions of the invention can be used in the following:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M. G. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, N.Y. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I. K. and Briddell, R. A. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, N.Y. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R. E. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 1-21, Wiley-Liss, Inc., New York, N.Y. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, N.Y. 1994; Long term culture initiating cell assay, Sutherland, H. J. In Culture of Hematopoietic Cells. R. I. Freshney, et al. eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, N.Y. 1994.

15

20

25

10

4.10.6 TISSUE GROWTH ACTIVITY

A polypeptide of the present invention also may be involved in bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as in wound healing and tissue repair and replacement, and in healing of burns, incisions and ulcers.

A polypeptide of the present invention which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Compositions of a polypeptide, antibody, binding partner, or other modulator of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. De novo bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A polypeptide of this invention may also be involved in attracting bone-forming cells, stimulating growth of bone-forming cells, or inducing differentiation of progenitors of bone-forming cells. Treatment of osteoporosis, osteoarthritis, bone degenerative disorders, or periodontal disease, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes may also be possible using the composition of the invention.

Another category of tissue regeneration activity that may involve the polypeptide of the present invention is tendon/ligament formation. Induction of tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors ex vivo for return in vivo to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

10

15

20

25

30

35

The compositions of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a composition may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a composition of the invention.

Compositions of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

Compositions of the present invention may also be involved in the generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine,

kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring may allow normal tissue to regenerate. A polypeptide of the present invention may also exhibit angiogenic activity.

A composition of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A composition of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

Therapeutic compositions of the invention can be used in the following:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, H. I. and Rovee, D. T., eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

20

25

30

15

5

10

4.10.7 IMMUNE STIMULATING OR SUPPRESSING ACTIVITY

A polypeptide of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A polynucleotide of the invention can encode a polypeptide exhibiting such activities. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpes viruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, proteins of the present invention may also be useful where a boost to the immune system generally may be desirable, i.e., in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome. autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein (or antagonists thereof. including antibodies) of the present invention may also to be useful in the treatment of allergic reactions and conditions (e.g., anaphylaxis, serum sickness, drug reactions, food allergies, insect venom allergies, mastocytosis, allergic rhinitis, hypersensitivity pneumonitis, urticaria, angioedema, eczema, atopic dermatitis, allergic contact dermatitis, erythema multiforme, Stevens-Johnson syndrome, allergic conjunctivitis, atopic keratoconjunctivitis, venereal keratoconjunctivitis, giant papillary conjunctivitis and contact allergies), such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for example, organ transplantation), may also be treatable using a protein (or antagonists thereof) of the present invention. The therapeutic effects of the polypeptides or antagonists thereof on allergic reactions can be evaluated by in vivo animals models such as the cumulative contact enhancement test (Lastborn et al., Toxicology 125: 59-66, 1998), skin prick test (Hoffmann et al., Allergy 54: 446-54, 1999), guinea pig skin sensitization test (Vohr et al., Arch. Toxocol. 73: 501-9), and murine local lymph node assay (Kimber et al., J. Toxicol. Environ. Health 53: 563-79).

10

20

25

30

35

Using the proteins of the invention it may also be possible to modulate immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as, for example, B7)), e.g., preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue

transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a therapeutic composition of the invention may prevent cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, a lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular therapeutic compositions in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins in vivo as described in Lenschow et al., Science 257:789-792 (1992) and Turka et al., Proc. Natl. Acad. Sci USA, 89:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of therapeutic compositions of the invention on the development of that disease.

10

15

20

25

30

35

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block stimulation of T cells can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/lpr/lpr mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (e.g., a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial

immune response. For example, enhancing an immune response may be useful in cases of viral infection, including systemic viral diseases such as influenza, the common cold, and encephalitis.

5

10

15

20

25

30

35

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells in vitro with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the in vitro activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells in vivo.

A polypeptide of the present invention may provide the necessary stimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient mounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (e.g., a cytoplasmic-domain truncated portion) of an MHC class I alpha chain protein and β_2 microglobulin protein or an MHC class II alpha chain protein and an MHC class II beta chain protein to thereby express MHC class I or MHC class II proteins on the cell surface. Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (e.g., B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J.

Immunol. 135:1564-1572, 1985; Takai et al., I. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bowman et al., J. Virology 61:1992-1998; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: In vitro antibody production, Mond, J. J. and Brunswick, M. In Current Protocols in Immunology. J. E. e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

10

15

20

25

30

35

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

4.10.8 ACTIVIN/INHIBIN ACTIVITY

10

20

25

30

35

A polypeptide of the present invention may also exhibit activin- or inhibin-related activities. A polynucleotide of the invention may encode a polypeptide exhibiting such characteristics. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a polypeptide of the present invention, alone or in heterodimers with a member of the inhibin family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the polypeptide of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, U.S. Pat. No. 4,798,885. A polypeptide of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as, but not limited to, cows, sheep and pigs.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods.

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

4.10.9 CHEMOTACTIC/CHEMOKINETIC ACTIVITY

A polypeptide of the present invention may be involved in chemotactic or chemokinetic activity for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Chemotactic and chemokinetic receptor activation can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic compositions (e.g. proteins, antibodies, binding partners, or modulators of the invention) provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

Therapeutic compositions of the invention can be used in the following:

5

10

15

20

25

35

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Marguiles, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25:1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153:1762-1768, 1994.

4.10.10 HEMOSTATIC AND THROMBOLYTIC ACTIVITY

A polypeptide of the invention may also be involved in hemostatis or thrombolysis or thrombosis. A polynucleotide of the invention can encode a polypeptide exhibiting such attributes. Compositions may be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A composition of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

Therapeutic compositions of the invention can be used in the following:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

4.10.11 CANCER DIAGNOSIS AND THERAPY

Polypeptides of the invention may be involved in cancer cell generation, proliferation or metastasis. Detection of the presence or amount of polynucleotides or polypeptides of the

invention may be useful for the diagnosis and/or prognosis of one or more types of cancer. For example, the presence or increased expression of a polynucleotide/polypeptide of the invention may indicate a hereditary risk of cancer, a precancerous condition, or an ongoing malignancy. Conversely, a defect in the gene or absence of the polypeptide may be associated with a cancer condition. Identification of single nucleotide polymorphisms associated with cancer or a predisposition to cancer may also be useful for diagnosis or prognosis.

5

10

15

20

25

30

35

Cancer treatments promote tumor regression by inhibiting tumor cell proliferation, inhibiting angiogenesis (growth of new blood vessels that is necessary to support tumor growth) and/or prohibiting metastasis by reducing tumor cell motility or invasiveness. Therapeutic compositions of the invention may be effective in adult and pediatric oncology including in solid phase tumors/malignancies, locally advanced tumors, human soft tissue sarcomas, metastatic cancer, including lymphatic metastases, blood cell malignancies including multiple myeloma. acute and chronic leukemias, and lymphomas, head and neck cancers including mouth cancer, larynx cancer and thyroid cancer, lung cancers including small cell carcinoma and non-small cell cancers, breast cancers including small cell carcinoma and ductal carcinoma, gastrointestinal cancers including esophageal cancer, stomach cancer, colon cancer, colorectal cancer and polyps associated with colorectal neoplasia, pancreatic cancers, liver cancer, urologic cancers including bladder cancer and prostate cancer, malignancies of the female genital tract including ovarian carcinoma, uterine (including endometrial) cancers, and solid tumor in the ovarian follicle, kidney cancers including renal cell carcinoma, brain cancers including intrinsic brain tumors, neuroblastoma, astrocytic brain tumors, gliomas, metastatic tumor cell invasion in the central nervous system, bone cancers including osteomas, skin cancers including malignant melanoma, tumor progression of human skin keratinocytes, squamous cell carcinoma, basal cell carcinoma, hemangiopericytoma and Karposi's sarcoma.

Polypeptides, polynucleotides, or modulators of polypeptides of the invention (including inhibitors and stimulators of the biological activity of the polypeptide of the invention) may be administered to treat cancer. Therapeutic compositions can be administered in therapeutically effective dosages alone or in combination with adjuvant cancer therapy such as surgery, chemotherapy, radiotherapy, thermotherapy, and laser therapy, and may provide a beneficial effect, e.g. reducing tumor size, slowing rate of tumor growth, inhibiting metastasis, or otherwise improving overall clinical condition, without necessarily eradicating the cancer.

The composition can also be administered in therapeutically effective amounts as a portion of an anti-cancer cocktail. An anti-cancer cocktail is a mixture of the polypeptide or modulator of the invention with one or more anti-cancer drugs in addition to a pharmaceutically acceptable carrier for delivery. The use of anti-cancer cocktails as a cancer treatment is routine.

Anti-cancer drugs that are well known in the art and can be used as a treatment in combination with the polypeptide or modulator of the invention include: Actinomycin D, Aminoglutethimide, Asparaginase, Bleomycin, Busulfan, Carboplatin, Carmustine, Chlorambucil, Cisplatin (cis-DDP), Cyclophosphamide, Cytarabine HCl (Cytosine arabinoside), Dacarbazine, Dactinomycin, Daunorubicin HCl, Doxorubicin HCl, Estramustine phosphate sodium, Etoposide (V16-213), Floxuridine, 5-Fluorouracil (5-Fu), Flutamide, Hydroxyurea (hydroxycarbamide), Ifosfamide, Interferon Alpha-2a, Interferon Alpha-2b, Leuprolide acetate (LHRH-releasing factor analog), Lomustine, Mechlorethamine HCl (nitrogen mustard), Melphalan, Mercaptopurine, Mesna, Methotrexate (MTX), Mitomycin, Mitoxantrone HCl, Octreotide, Plicamycin, Procarbazine HCl, Streptozocin, Tamoxifen citrate, Thioguanine, Thiotepa, Vinblastine sulfate, Vincristine sulfate, Amsacrine, Azacitidine, Hexamethylmelamine, Interleukin-2, Mitoguazone, Pentostatin, Semustine, Teniposide, and Vindesine sulfate.

5

10

15

20

25

35

In addition, therapeutic compositions of the invention may be used for prophylactic treatment of cancer. There are hereditary conditions and/or environmental situations (e.g. exposure to carcinogens) known in the art that predispose an individual to developing cancers. Under these circumstances, it may be beneficial to treat these individuals with therapeutically effective doses of the polypeptide of the invention to reduce the risk of developing cancers.

In vitro models can be used to determine the effective doses of the polypeptide of the invention as a potential cancer treatment. These in vitro models include proliferation assays of cultured tumor cells, growth of cultured tumor cells in soft agar (see Freshney, (1987) Culture of Animal Cells: A Manual of Basic Technique, Wily-Liss, New York, NY Ch 18 and Ch 21), tumor systems in nude mice as described in Giovanella et al., J. Natl. Can. Inst., 52: 921-30 (1974), mobility and invasive potential of tumor cells in Boyden Chamber assays as described in Pilkington et al., Anticancer Res., 17: 4107-9 (1997), and angiogenesis assays such as induction of vascularization of the chick chorioallantoic membrane or induction of vascular endothelial cell migration as described in Ribatta et al., Intl. J. Dev. Biol., 40: 1189-97 (1999) and Li et al., Clin. Exp. Metastasis, 17:423-9 (1999), respectively. Suitable tumor cells lines are available, e.g. from American Type Tissuc Culture Collection catalogs.

30 4.10.12 RECEPTOR/LIGAND ACTIVITY

A polypeptide of the present invention may also demonstrate activity as receptor, receptor ligand or inhibitor or agonist of receptor/ligand interactions. A polynucleotide of the invention can encode a polypeptide exhibiting such characteristics. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions

and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses. Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a polypeptide of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A. M. Kruisbeek, D. H. Margulies, E. M. Shevach, W. Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

By way of example, the polypeptides of the invention may be used as a receptor for a ligand(s) thereby transmitting the biological activity of that ligand(s). Ligands may be identified through binding assays, affinity chromatography, dihybrid screening assays, BIAcore assays, gel overlay assays, or other methods known in the art.

Studies characterizing drugs or proteins as agonist or antagonist or partial agonists or a partial antagonist require the use of other proteins as competing ligands. The polypeptides of the present invention or ligand(s) thereof may be labeled by being coupled to radioisotopes, colorimetric molecules or a toxin molecules by conventional methods. ("Guide to Protein Purification" Murray P. Deutscher (ed) Methods in Enzymology Vol. 182 (1990) Academic Press, Inc. San Diego). Examples of radioisotopes include, but are not limited to, tritium and carbon-14. Examples of colorimetric molecules include, but are not limited to, fluorescent molecules such as fluorescamine, or rhodamine or other colorimetric molecules. Examples of toxins include, but are not limited, to ricin.

4.10.13 DRUG SCREENING

10

15

20

25

30

35

This invention is particularly useful for screening chemical compounds by using the novel polypeptides or binding fragments thereof in any of a variety of drug screening techniques. The polypeptides or fragments employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface or located intracellularly. One method of drug screening

utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or a fragment thereof. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between polypeptides of the invention or fragments and the agent being tested or examine the diminution in complex formation between the novel polypeptides and an appropriate cell line, which are well known in the art.

Sources for test compounds that may be screened for ability to bind to or modulate (i.e., increase or decrease) the activity of polypeptides of the invention include (1) inorganic and organic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of either random or mimetic peptides, oligonucleotides or organic molecules.

10

15

20

25

30

35

Chemical libraries may be readily synthesized or purchased from a number of commercial sources, and may include structural analogs of known compounds or compounds that are identified as "hits" or "leads" via natural product screening.

The sources of natural product libraries are microorganisms (including bacteria and fungi), animals, plants or other vegetation, or marine organisms, and libraries of mixtures for screening may be created by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of the organisms themselves. Natural product libraries include polyketides, non-ribosomal peptides, and (non-naturally occurring) variants thereof. For a review, see *Science* 282:63-68 (1998).

Combinatorial libraries are composed of large numbers of peptides, oligonucleotides or organic compounds and can be readily prepared by traditional automated synthesis methods, PCR, cloning or proprietary synthetic methods. Of particular interest are peptide and oligonucleotide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, Curr. Opin. Biotechnol. 8:701-707 (1997). For reviews and examples of peptidomimetic libraries, see Al-Obeidi et al., Mol. Biotechnol, 9(3):205-23 (1998); Hruby et al., Curr Opin Chem Biol, 1(1):114-19 (1997); Dorner et al., Bioorg Med Chem, 4(5):709-15 (1996) (alkylated dipeptides).

Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to bind a polypeptide of the invention. The molecules identified in the binding assay are then tested for antagonist or agonist activity in *in vivo* tissue culture or animal models that are well known in the art. In brief, the molecules are titrated into a plurality of cell cultures or animals and then tested for either cell/animal death or prolonged survival of the animal/cells.

The binding molecules thus identified may be complexed with toxins, e.g., ricin or cholera, or with other compounds that are toxic to cells such as radioisotopes. The toxin-binding molecule complex is then targeted to a tumor or other cell by the specificity of the binding molecule for a polypeptide of the invention. Alternatively, the binding molecules may be complexed with imaging agents for targeting and imaging purposes.

4.10.14 ASSAY FOR RECEPTOR ACTIVITY

The invention also provides methods to detect specific binding of a polypeptide e.g. a ligand or a receptor. The art provides numerous assays particularly useful for identifying previously unknown binding partners for receptor polypeptides of the invention. For example, expression cloning using mammalian or bacterial cells, or dihybrid screening assays can be used to identify polynucleotides encoding binding partners. As another example, affinity chromatography with the appropriate immobilized polypeptide of the invention can be used to isolate polypeptides that recognize and bind polypeptides of the invention. There are a number of different libraries used for the identification of compounds, and in particular small molecules. that modulate (i.e., increase or decrease) biological activity of a polypeptide of the invention. Ligands for receptor polypeptides of the invention can also be identified by adding exogenous ligands, or cocktails of ligands to two cells populations that are genetically identical except for the expression of the receptor of the invention: one cell population expresses the receptor of the invention whereas the other does not. The response of the two cell populations to the addition of ligands(s) are then compared. Alternatively, an expression library can be co-expressed with the polypeptide of the invention in cells and assayed for an autocrine response to identify potential ligand(s). As still another example, BIAcore assays, gel overlay assays, or other methods known in the art can be used to identify binding partner polypeptides, including, (1) organic and inorganic chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules.

The role of downstream intracellular signaling molecules in the signaling cascade of the polypeptide of the invention can be determined. For example, a chimeric protein in which the cytoplasmic domain of the polypeptide of the invention is fused to the extracellular portion of a protein, whose ligand has been identified, is produced in a host cell. The cell is then incubated with the ligand specific for the extracellular portion of the chimeric protein, thereby activating the chimeric receptor. Known downstream proteins involved in intracellular signaling can then be assayed for expected modifications *i.e.* phosphorylation. Other methods known to those in the art can also be used to identify signaling molecules involved in receptor activity.

5

10

15

20

25

30

4.10.15 ANTI-INFLAMMATORY ACTIVITY

Compositions of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Compositions with such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation intimation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Compositions of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material. Compositions of this invention may be utilized to prevent or treat conditions such as, but not limited to, sepsis, acute pancreatitis, endotoxin shock, cytokine induced shock, rheumatoid arthritis, chronic inflammatory arthritis, pancreatic cell damage from diabetes mellitus type 1, graft versus host disease, inflammatory bowel disease, inflamation associated with pulmonary disease, other autoimmune disease or inflammatory disease, an antiproliferative agent such as for acute or chronic mylegenous leukemia or in the prevention of premature labor secondary to intrauterine infections.

4.10.16 LEUKEMIAS

10

15

20

35

Leukemias and related disorders may be treated or prevented by administration of a

therapeutic that promotes or inhibits function of the polynucleotides and/or polypeptides of the invention. Such leukemias and related disorders include but are not limited to acute leukemia, acute lymphocytic leukemia, acute myelocytic leukemia, myeloblastic, promyelocytic, myelomonocytic, monocytic, erythroleukemia, chronic leukemia, chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia (for a review of such disorders, see

Fishman et al., 1985, Medicine, 2d Ed., J.B. Lippincott Co., Philadelphia).

4.10.17 NERVOUS SYSTEM DISORDERS

Nervous system disorders, involving cell types which can be tested for efficacy of intervention with compounds that modulate the activity of the polynucleotides and/or polypeptides of the invention, and which can be treated upon thus observing an indication of

therapeutic utility, include but are not limited to nervous system injuries, and diseases or disorders which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be treated in a patient (including human and non-human mammalian patients) according to the invention include but are not limited to the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems:

- (i) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries;
- 10 (ii) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia;

15

20

25

30

- (iii) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis;
- (iv) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis;
- (v) lesions associated with nutritional diseases or disorders, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration;
- (vi) neurological lesions associated with systemic diseases including but not limited to diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis;
- (vii) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and
- (viii) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including but not limited to multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

Therapeutics which are useful according to the invention for treatment of a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, therapeutics which elicit any of the following effects may be useful according to the invention:

(i) increased survival time of neurons in culture;

5

10

15

20

25

30

35

- (ii) increased sprouting of neurons in culture or in vivo;
- (iii) increased production of a neuron-associated molecule in culture or *in vivo*, *e.g.*, choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or
 - (iv) decreased symptoms of neuron dysfunction in vivo.

Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may be measured by the method set forth in Arakawa et al. (1990, J. Neurosci. 10:3507-3515); increased sprouting of neurons may be detected by methods set forth in Pestronk et al. (1980, Exp. Neurol. 70:65-82) or Brown et al. (1981, Ann. Rev. Neurosci. 4:17-42); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, etc., depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

In specific embodiments, motor neuron disorders that may be treated according to the invention include but are not limited to disorders such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as disorders that selectively affect neurons such as amyotrophic lateral sclerosis, and including but not limited to progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

4.10.18 OTHER ACTIVITIES

A polypeptide of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape);

effecting biorhythms or circadian cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, co-factors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

4.10.19 IDENTIFICATION OF POLYMORPHISMS

10

15

20

25

30

35

The demonstration of polymorphisms makes possible the identification of such polymorphisms in human subjects and the pharmacogenetic use of this information for diagnosis and treatment. Such polymorphisms may be associated with, e.g., differential predisposition or susceptibility to various disease states (such as disorders involving inflammation or immune response) or a differential response to drug administration, and this genetic information can be used to tailor preventive or therapeutic treatment appropriately. For example, the existence of a polymorphism associated with a predisposition to inflammation or autoimmune disease makes possible the diagnosis of this condition in humans by identifying the presence of the polymorphism.

Polymorphisms can be identified in a variety of ways known in the art which all generally involve obtaining a sample from a patient, analyzing DNA from the sample, optionally involving isolation or amplification of the DNA, and identifying the presence of the polymorphism in the DNA. For example, PCR may be used to amplify an appropriate fragment of genomic DNA which may then be sequenced. Alternatively, the DNA may be subjected to allele-specific oligonucleotide hybridization (in which appropriate oligonucleotides are hybridized to the DNA under conditions permitting detection of a single base mismatch) or to a single nucleotide extension assay (in which an oligonucleotide that hybridizes immediately adjacent to the position of the polymorphism is extended with one or more labeled nucleotides). In addition, traditional restriction fragment length polymorphism analysis (using restriction enzymes that provide differential digestion of the genomic DNA depending on the presence or

absence of the polymorphism) may be performed. Arrays with nucleotide sequences of the present invention can be used to detect polymorphisms. The array can comprise modified nucleotide sequences of the present invention in order to detect the nucleotide sequences of the present invention. In the alternative, any one of the nucleotide sequences of the present invention can be placed on the array to detect changes from those sequences.

Alternatively a polymorphism resulting in a change in the amino acid sequence could also be detected by detecting a corresponding change in amino acid sequence of the protein, e.g., by an antibody specific to the variant sequence.

10 4.10.20 ARTHRITIS AND INFLAMMATION

The immunosuppressive effects of the compositions of the invention against rheumatoid arthritis is determined in an experimental animal model system. The experimental model system is adjuvant induced arthritis in rats, and the protocol is described by J. Holoshitz, et at., 1983, Science, 219:56, or by B. Waksman et al., 1963, Int. Arch. Allergy Appl. Immunol., 23:129. Induction of the disease can be caused by a single injection, generally intradermally, of a suspension of killed Mycobacterium tuberculosis in complete Freund's adjuvant (CFA). The route of injection can vary, but rats may be injected at the base of the tail with an adjuvant mixture. The polypeptide is administered in phosphate buffered solution (PBS) at a dose of about 1-5 mg/kg. The control consists of administering PBS only.

The procedure for testing the effects of the test compound would consist of intradermally injecting killed Mycobacterium tuberculosis in CFA followed by immediately administering the test compound and subsequent treatment every other day until day 24. At 14, 15, 18, 20, 22, and 24 days after injection of Mycobacterium CFA, an overall arthritis score may be obtained as described by J. Holoskitz above. An analysis of the data would reveal that the test compound would have a dramatic affect on the swelling of the joints as measured by a decrease of the arthritis score.

4.11 THERAPEUTIC METHODS

15

20

25

30

The compositions (including polypeptide fragments, analogs, variants and antibodies or other binding partners or modulators including antisense polynucleotides) of the invention have numerous applications in a variety of therapeutic methods. Examples of therapeutic applications include, but are not limited to, those exemplified herein.

4.11.1 EXAMPLE

One embodiment of the invention is the administration of an effective amount of the polypeptides or other composition of the invention to individuals affected by a disease or disorder that can be modulated by regulating the peptides of the invention. While the mode of administration is not particularly important, parenteral administration is preferred. An exemplary mode of administration is to deliver an intravenous bolus. The dosage of the polypeptides or other composition of the invention will normally be determined by the prescribing physician. It is to be expected that the dosage will vary according to the age, weight, condition and response of the individual patient. Typically, the amount of polypeptide administered per dose will be in the range of about 0.01µg/kg to 100 mg/kg of body weight, with the preferred dose being about 0.1µg/kg to 10 mg/kg of patient body weight. For parenteral administration, polypeptides of the invention will be formulated in an injectable form combined with a pharmaceutically acceptable parenteral vehicle. Such vehicles are well known in the art and examples include water, saline, Ringer's solution, dextrose solution, and solutions consisting of small amounts of the human serum albumin. The vehicle may contain minor amounts of additives that maintain the isotonicity and stability of the polypeptide or other active ingredient. The preparation of such solutions is within the skill of the art.

4.12 PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

10

15

20

25

30

35

A protein or other composition of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources and including antibodies and other binding partners of the polypeptides of the invention) may be administered to a patient in need, by itself, or in pharmaceutical compositions where it is mixed with suitable carriers or excipient(s) at doses to treat or ameliorate a variety of disorders. Such a composition may optionally contain (in addition to protein or other active ingredient and a carrier) diluents. fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the disease or disorder in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet-derived growth

factor (PDGF), transforming growth factors (TGF- α and TGF- β), insulin-like growth factor (IGF), as well as cytokines described herein.

The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or other active ingredient or complement its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein or other active ingredient of the invention, or to minimize side effects. Conversely, protein or other active ingredient of the present invention may be included in formulations of the particular clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the clotting factor, cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent (such as IL-1Ra, IL-1 Hy1, IL-1 Hy2, anti-TNF, corticosteroids, immunosuppressive agents). A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

As an alternative to being included in a pharmaceutical composition of the invention including a first protein, a second protein or a therapeutic agent may be concurrently administered with the first protein (e.g., at the same time, or at differing times provided that therapeutic concentrations of the combination of agents is achieved at the treatment site). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition. A therapeutically effective dose further refers to that amount of the compound sufficient to result in amelioration of symptoms, e.g., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, a therapeutically effective dose refers to that ingredient alone. When applied to a combination, a therapeutically effective dose refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein or other active ingredient of the present invention is administered to a mammal having a condition to be treated. Protein or other active ingredient of the present invention may be administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co- administered with one or more cytokines, lymphokines or other

hematopoietic factors, protein or other active ingredient of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein or other active ingredient of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

4.12.1 ROUTES OF ADMINISTRATION

10

20

25

30

35

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. Administration of protein or other active ingredient of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a arthritic joints or in fibrotic tissue, often in a depot or sustained release formulation. In order to prevent the scarring process frequently occurring as complication of glaucoma surgery, the compounds may be administered topically, for example, as eye drops. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with a specific antibody, targeting, for example, arthritic or fibrotic tissue. The liposomes will be targeted to and taken up selectively by the afflicted tissue.

The polypeptides of the invention are administered by any route that delivers an effective dosage to the desired site of action. The determination of a suitable route of administration and an effective dosage for a particular indication is within the level of skill in the art. Preferably for wound treatment, one administers the therapeutic compound directly to the site. Suitable dosage ranges for the polypeptides of the invention can be extrapolated from these dosages or from similar studies in appropriate animal models. Dosages can then be adjusted as necessary by the clinician to provide maximal therapeutic benefit.

4.12.2 COMPOSITIONS/FORMULATIONS

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in a conventional manner using one or more physiologically acceptable carriers

comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. These pharmaceutical compositions may be manufactured in a manner that is itself known, e.g., by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes. Proper formulation is dependent upon the route of administration chosen. When a therapeutically effective amount of protein or other active ingredient of the present invention is administered orally, protein or other active ingredient of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein or other active ingredient of the present invention, and preferably from about 25 to 90% protein or other active ingredient of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein or other active ingredient of the present invention, and preferably from about 1 to 50% protein or other active ingredient of the present invention.

10

15

20

25

30

35

When a therapeutically effective amount of protein or other active ingredient of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein or other active ingredient of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein or other active ingredient solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein or other active ingredient of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art. For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Pharmaceutical preparations for oral use can be obtained from a solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, tale, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

10

15

20

25

30

35

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration. For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, e.g., dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of, e.g., gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch. The compounds may be formulated for parenteral

administration by injection, e.g., by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, e.g., in ampules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions. Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, e.g., sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, e.g., containing conventional suppository bases such as cocoa butter or other glycerides. In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a co-solvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The co-solvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may replace polyethylene glycol, e.g. polyvinyl pyrrolidone; and other

sugars or polysaccharides may substitute for dextrose. Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity.

Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various types of sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein or other active ingredient stabilization may be employed.

10

15

20

25

30

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols. Many of the active ingredients of the invention may be provided as salts with pharmaceutically compatible counter ions. Such pharmaceutically acceptable base addition salts are those salts which retain the biological effectiveness and properties of the free acids and which are obtained by reaction with inorganic or organic bases such as sodium hydroxide, magnesium hydroxide, ammonia, trialkylamine, dialkylamine, monoalkylamine, dibasic amino acids, sodium acetate, potassium benzoate, triethanol amine and the like.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) or other active ingredient(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunoglobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically

acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithins, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent Nos. 4,235,871; 4,501,728; 4,837,028; and 4,737,323, all of which are incorporated herein by reference.

10

15

20

25

30

The amount of protein or other active ingredient of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein or other active ingredient of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein or other active ingredient of the present invention and observe the patient's response. Larger doses of protein or other active ingredient of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1 µg to about 10 mg, more preferably about 0.1 ug to about 1 mg) of protein or other active ingredient of the present invention per kg body weight. For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein or other active ingredient of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing or other active ingredient-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalcium phosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxyapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalcium phosphate. The bioceramics may be altered in composition, such as in calcium-aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability. Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

10

20

25

30

35

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt %, preferably 1-10 wt % based on total formulation weight, which represents the amount necessary to prevent desorption of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells. In further compositions, proteins or other active ingredients of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications.

Particularly domestic animals and thoroughbred horses, in addition to humans, are desired

patients for such treatment with proteins or other active ingredients of the present invention. The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, *e.g.*, amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (*e.g.*, bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either in vivo or ex vivo into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA). Cells may also be cultured ex vivo in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced in vivo for therapeutic purposes.

20

25

30

10

15

4.12.3 EFFECTIVE DOSAGE

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount effective to prevent development of or to alleviate the existing symptoms of the subject being treated. Determination of the effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein. For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from appropriate in vitro assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that can be used to more accurately determine useful doses in humans. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC₅₀ as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the protein's biological activity). Such information can be used to more accurately determine useful doses in humans.

A therapeutically effective dose refers to that amount of the compound that results in amelioration of symptoms or a prolongation of survival in a patient. Toxicity and therapeutic efficacy of such compounds can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, e.g., for determining the LD₅₀ (the dose lethal to 50% of the population) and the ED₅₀ (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD₅₀ and ED₅₀. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED₅₀ with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. See, e.g., Fingl et al., 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1. Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the desired effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from in vitro data. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%. In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

An exemplary dosage regimen for polypeptides or other compositions of the invention will be in the range of about $0.01~\mu g/kg$ to 100~mg/kg of body weight daily, with the preferred dose being about $0.1~\mu g/kg$ to 25~mg/kg of patient body weight daily, varying in adults and children. Dosing may be once daily, or equivalent doses may be delivered at longer or shorter intervals.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's age and weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

4.12.4 PACKAGING

10

15

20

25

30

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may, for example, comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition.

4.13 ANTIBODIES

5

10

15

20

25

30

35

Also included in the invention are antibodies to proteins, or fragments of proteins of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin (Ig) molecules, *i.e.*, molecules that contain an antigen binding site that specifically binds (immunoreacts with) an antigen. Such antibodies include, but are not limited to, polyclonal, monoclonal, chimeric, single chain, F_{ab} , F_{ab} and $F_{(ab)}$ fragments, and an F_{ab} expression library. In general, an antibody molecule obtained from humans relates to any of the classes IgG, IgM, IgA, IgE and IgD, which differ from one another by the nature of the heavy chain present in the molecule. Certain classes have subclasses as well, such as IgG₁, IgG₂, and others. Furthermore, in humans, the light chain may be a kappa chain or a lambda chain. Reference herein to antibodies includes a reference to all such classes, subclasses and types of human antibody species.

An isolated related protein of the invention may be intended to serve as an antigen, or a portion or fragment thereof, and additionally can be used as an immunogen to generate antibodies that immunospecifically bind the antigen, using standard techniques for polyclonal and monoclonal antibody preparation. The full-length protein can be used or, alternatively, the invention provides antigenic peptide fragments of the antigen for use as immunogens. An antigenic peptide fragment comprises at least 6 amino acid residues of the amino acid sequence of the full length protein, (for example the amino acid sequence shown in SEQ ID NO: 1351), and encompasses an epitope thereof such that an antibody raised against the peptide forms a specific immune complex with the full length protein or with any fragment that contains the epitope. Preferably, the antigenic peptide comprises at least 10 amino acid residues, or at least 15 amino acid residues, or at least 20 amino acid residues. Preferred epitopes encompassed by the antigenic peptide are regions of the protein that are located on its surface; commonly these are hydrophilic regions.

In certain embodiments of the invention, at least one epitope encompassed by the antigenic peptide is a region of -related protein that is located on the surface of the protein, e.g., a hydrophilic region. A hydrophobicity analysis of the human related protein sequence will

indicate which regions of a related protein are particularly hydrophilic and, therefore, are likely to encode surface residues useful for targeting antibody production. As a means for targeting antibody production, hydropathy plots showing regions of hydrophilicity and hydrophobicity may be generated by any method well known in the art, including, for example, the Kyte Doolittle or the Hopp Woods methods, either with or without Fourier transformation. See, e.g., Hopp and Woods, 1981, *Proc. Nat. Acad. Sci. USA* 78: 3824-3828; Kyte and Doolittle 1982, J. Mol. Biol. 157: 105-142, each of which is incorporated herein by reference in its entirety. Antibodies that are specific for one or more domains within an antigenic protein, or derivatives, fragments, analogs or homologs thereof, are also provided herein.

A protein of the invention, or a derivative, fragment, analog, homolog or ortholog thereof, may be utilized as an immunogen in the generation of antibodies that immunospecifically bind these protein components.

Various procedures known within the art may be used for the production of polyclonal or monoclonal antibodies directed against a protein of the invention, or against derivatives, fragments, analogs homologs or orthologs thereof (see, for example, Antibodies: A Laboratory Manual, Harlow E, and Lane D, 1988, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, incorporated herein by reference). Some of these antibodies are discussed below.

5.13.1 Polyclonal Antibodies

10

15

20

25

30

For the production of polyclonal antibodies, various suitable host animals (e.g., rabbit, goat, mouse or other mammal) may be immunized by one or more injections with the native protein, a synthetic variant thereof, or a derivative of the foregoing. An appropriate immunogenic preparation can contain, for example, the naturally occurring immunogenic protein, a chemically synthesized polypeptide representing the immunogenic protein, or a recombinantly expressed immunogenic protein. Furthermore, the protein may be conjugated to a second protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. The preparation can further include an adjuvant. Various adjuvants used to increase the immunological response include, but are not limited to, Freund's (complete and incomplete), mineral gels (e.g., aluminum hydroxide), surface active substances (e.g., lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, dinitrophenol, etc.), adjuvants usable in humans such as Bacille Calmette-Guerin and Corynebacterium parvum, or similar immunostimulatory agents. Additional examples of adjuvants which can be employed include MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).

The polyclonal antibody molecules directed against the immunogenic protein can be isolated from the mammal (e.g., from the blood) and further purified by well known techniques, such as affinity chromatography using protein A or protein G, which provide primarily the IgG fraction of immune serum. Subsequently, or alternatively, the specific antigen which is the target of the immunoglobulin sought, or an epitope thereof, may be immobilized on a column to purify the immune specific antibody by immunoaffinity chromatography. Purification of immunoglobulins is discussed, for example, by D. Wilkinson (The Scientist, published by The Scientist, Inc., Philadelphia PA, Vol. 14, No. 8 (April 17, 2000), pp. 25-28).

10 5.13.2 Monoclonal Antibodies

15

20

25

30

35

The term "monoclonal antibody" (MAb) or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one molecular species of antibody molecule consisting of a unique light chain gene product and a unique heavy chain gene product. In particular, the complementarity determining regions (CDRs) of the monoclonal antibody are identical in all the molecules of the population. MAbs thus contain an antigen binding site capable of immunoreacting with a particular epitope of the antigen characterized by a unique binding affinity for it.

Monoclonal antibodies can be prepared using hybridoma methods, such as those described by Kohler and Milstein, <u>Nature</u>, <u>256</u>:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes can be immunized in vitro.

The immunizing agent will typically include the protein antigen, a fragment thereof or a fusion protein thereof. Generally, either peripheral blood lymphocytes are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells can be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, <u>J. Immunol.</u>, <u>133</u>:3001 (1984); Brodeur et al., <u>Monoclonal Antibody Production Techniques and Applications</u>, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the antigen. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, <u>Anal. Biochem.</u>, <u>107</u>:220 (1980). Preferably, antibodies having a high degree of specificity and a high binding affinity for the target antigen are isolated.

10

15

20

25

30

35

After the desired hybridoma cells are identified, the clones can be subcloned by limiting dilution procedures and grown by standard methods. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium.

Alternatively, the hybridoma cells can be grown in vivo as ascites in a mammal.

The monoclonal antibodies secreted by the subclones can be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies can also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA can be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also can be modified, for

example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Patent No. 4,816,567; Morrison, Nature 368, 812-13 (1994)) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

5.13.2 Humanized Antibodies

10

15

20

25

30

35

The antibodies directed against the protein antigens of the invention can further comprise humanized antibodies or human antibodies. These antibodies are suitable for administration to humans without engendering an immune response by the human against the administered immunoglobulin. Humanized forms of antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')2 or other antigenbinding subsequences of antibodies) that are principally comprised of the sequence of a human immunoglobulin, and contain minimal sequence derived from a non-human immunoglobulin. Humanization can be performed following the method of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988)), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. (See also U.S. Patent No. 5,225,539.) In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies can also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the framework regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., 1986; Riechmann et al., 1988; and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)).

5.13.3 Human Antibodies

Fully human antibodies relate to antibody molecules in which essentially the entire sequences of both the light chain and the heavy chain, including the CDRs, arise from human genes. Such antibodies are termed "human antibodies", or "fully human antibodies" herein.

Human monoclonal antibodies can be prepared by the trioma technique; the human B-cell hybridoma technique (see Kozbor, et al., 1983 Immunol Today 4: 72) and the EBV hybridoma technique to produce human monoclonal antibodies (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96). Human monoclonal antibodies may be utilized in the practice of the present invention and may be produced by using human hybridomas (see Cote, et al., 1983. Proc Natl Acad Sci USA 80: 2026-2030) or by transforming human B-cells with Epstein Barr Virus in vitro (see Cole, et al., 1985 In: MONOCLONAL ANTIBODIES AND CANCER THERAPY, Alan R. Liss, Inc., pp. 77-96).

In addition, human antibodies can also be produced using additional techniques, including phage display libraries (Hoogenboom and Winter, J. Mol. Biol., 227:381 (1991); Marks et al., J. Mol. Biol., 222:581 (1991)). Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in Marks et al. (Bio/Technology 10, 779-783 (1992)); Lonberg et al. (Nature 368 856-859 (1994)); Morrison (Nature 368, 812-13 (1994)); Fishwild et al., (Nature Biotechnology 14, 845-51 (1996)); Neuberger (Nature Biotechnology 14, 826 (1996)); and Lonberg and Huszar (Intern. Rev. Immunol. 13 65-93 (1995)).

Human antibodies may additionally be produced using transgenic nonhuman animals which are modified so as to produce fully human antibodies rather than the animal's endogenous antibodies in response to challenge by an antigen. (See PCT publication WO94/02602). The endogenous genes encoding the heavy and light immunoglobulin chains in the nonhuman host have been incapacitated, and active loci encoding human heavy and light chain immunoglobulins are inserted into the host's genome. The human genes are incorporated, for example, using yeast artificial chromosomes containing the requisite human DNA segments. An animal which provides all the desired modifications is then obtained as progeny by crossbreeding intermediate transgenic animals containing fewer than the full complement of the modifications. The preferred embodiment of such a nonhuman animal is a mouse, and is termed the XenomouseTM as disclosed in PCT publications WO 96/33735 and WO 96/34096. This animal produces B cells which secrete fully human immunoglobulins. The antibodies can be obtained directly from the animal after immunization with an immunogen of interest, as, for example, a preparation of a polyclonal antibody, or alternatively from immortalized B cells derived from the animal, such as hybridomas producing monoclonal antibodies. Additionally, the genes encoding the

immunoglobulins with human variable regions can be recovered and expressed to obtain the antibodies directly, or can be further modified to obtain analogs of antibodies such as, for example, single chain Fv molecules.

An example of a method of producing a nonhuman host, exemplified as a mouse, lacking expression of an endogenous immunoglobulin heavy chain is disclosed in U.S. Patent No. 5,939,598. It can be obtained by a method including deleting the J segment genes from at least one endogenous heavy chain locus in an embryonic stem cell to prevent rearrangement of the locus and to prevent formation of a transcript of a rearranged immunoglobulin heavy chain locus, the deletion being effected by a targeting vector containing a gene encoding a selectable marker; and producing from the embryonic stem cell a transgenic mouse whose somatic and germ cells contain the gene encoding the selectable marker.

A method for producing an antibody of interest, such as a human antibody, is disclosed in U.S. Patent No. 5,916,771. It includes introducing an expression vector that contains a nucleotide sequence encoding a heavy chain into one mammalian host cell in culture, introducing an expression vector containing a nucleotide sequence encoding a light chain into another mammalian host cell, and fusing the two cells to form a hybrid cell. The hybrid cell expresses an antibody containing the heavy chain and the light chain.

In a further improvement on this procedure, a method for identifying a clinically relevant epitope on an immunogen, and a correlative method for selecting an antibody that binds immunospecifically to the relevant epitope with high affinity, are disclosed in PCT publication WO 99/53049.

5.13.4 Fab Fragments and Single Chain Antibodies

According to the invention, techniques can be adapted for the production of single-chain antibodies specific to an antigenic protein of the invention (see e.g., U.S. Patent No. 4,946,778). In addition, methods can be adapted for the construction of F_{ab} expression libraries (see e.g., Huse, et al., 1989 Science 246: 1275-1281) to allow rapid and effective identification of monoclonal F_{ab} fragments with the desired specificity for a protein or derivatives, fragments, analogs or homologs thereof. Antibody fragments that contain the idiotypes to a protein antigen may be produced by techniques known in the art including, but not limited to: (i) an $F_{(ab)2}$ fragment produced by pepsin digestion of an antibody molecule; (ii) an F_{ab} fragment generated by reducing the disulfide bridges of an $F_{(ab)2}$ fragment; (iii) an F_{ab} fragment generated by the treatment of the antibody molecule with papain and a reducing agent and (iv) F_v fragments.

5.13.5 Bispecific Antibodies

5

10

15

20

25

30

35

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for an antigenic protein of the invention. The second binding target is any other antigen, and advantageously is a cell-surface protein or receptor or receptor subunit.

5

10

15

20

25

30

35

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983)). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker *et al.*, 1991 *EMBO J.*, 10:3655-3659.

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')₂ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., Science 229:81 (1985) describe a procedure

wherein intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Additionally, Fab' fragments can be directly recovered from E. coli and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from E. coli and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

10

15

20

25

30

35

Various techniques for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al., <u>J. Immunol.</u> 147:60 (1991). Exemplary bispecific antibodies can bind to two different epitopes, at least one of which originates in the protein antigen of the invention. Alternatively, an anti-antigenic arm of an immunoglobulin molecule can be combined with an arm which binds to a triggering molecule on

a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc R), such as Fc RI (CD64), Fc RII (CD32) and Fc RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular antigen. Bispecific antibodies can also be used to direct cytotoxic agents to cells which express a particular antigen. These antibodies possess an antigen-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the protein antigen described herein and further binds tissue factor (TF).

5.13.6 Heteroconjugate Antibodies

5

10

15

25

30

35

Heteroconjugate antibodies are also within the scope of the present invention.

Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Patent No. 4,676,980), and for treatment of HIV infection (WO 91/00360; WO 92/200373; EP 03089). It is contemplated that the antibodies can be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins can be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

20 5.13.7 Effector Function Engineering

It can be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) can be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated can have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity can also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and can thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

5.13.8 Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of

bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from Pseudomonas aeruginosa), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, Aleurites fordii proteins, dianthin proteins, Phytolaca americana proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ²¹²Bi, ¹³¹I, ¹³¹In, ⁹⁰Y, and ¹⁸⁶Re.

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutareldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody can be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is in turn conjugated to a cytotoxic agent.

4.14 COMPUTER READABLE SEQUENCES

5

10

15

20

25

30

35

In one application of this embodiment, a nucleotide sequence of the present invention can be recorded on computer readable media. As used herein, "computer readable media" refers to any medium which can be read and accessed directly by a computer. Such media include, but are not limited to: magnetic storage media, such as floppy discs, hard disc storage medium, and magnetic tape; optical storage media such as CD-ROM; electrical storage media such as RAM and ROM; and hybrids of these categories such as magnetic/optical storage media. A skilled

artisan can readily appreciate how any of the presently known computer readable mediums can be used to create a manufacture comprising computer readable medium having recorded thereon a nucleotide sequence of the present invention. As used herein, "recorded" refers to a process for storing information on computer readable medium. A skilled artisan can readily adopt any of the presently known methods for recording information on computer readable medium to generate manufactures comprising the nucleotide sequence information of the present invention.

5

15

20

30

35

A variety of data storage structures are available to a skilled artisan for creating a computer readable medium having recorded thereon a nucleotide sequence of the present invention. The choice of the data storage structure will generally be based on the means chosen to access the stored information. In addition, a variety of data processor programs and formats can be used to store the nucleotide sequence information of the present invention on computer readable medium. The sequence information can be represented in a word processing text file, formatted in commercially-available software such as WordPerfect and Microsoft Word, or represented in the form of an ASCII file, stored in a database application, such as DB2, Sybase, Oracle, or the like. A skilled artisan can readily adapt any number of data processor structuring formats (e.g. text file or database) in order to obtain computer readable medium having recorded thereon the nucleotide sequence information of the present invention.

By providing any of the nucleotide sequences SEQ ID NO:1-1350 or a representative fragment thereof; or a nucleotide sequence at least 95% identical to any of the nucleotide sequences of SEQ ID NO:1-1350 in computer readable form, a skilled artisan can routinely access the sequence information for a variety of purposes. Computer software is publicly available which allows a skilled artisan to access sequence information provided in a computer readable medium. The examples which follow demonstrate how software which implements the BLAST (Altschul et al., J. Mol. Biol. 215:403-410 (1990)) and BLAZE (Brutlag et al., Comp. Chem. 17:203-207 (1993)) search algorithms on a Sybase system is used to identify open reading frames (ORFs) within a nucleic acid sequence. Such ORFs may be protein encoding fragments and may be useful in producing commercially important proteins such as enzymes used in fermentation reactions and in the production of commercially useful metabolites.

As used herein, "a computer-based system" refers to the hardware means, software means, and data storage means used to analyze the nucleotide sequence information of the present invention. The minimum hardware means of the computer-based systems of the present invention comprises a central processing unit (CPU), input means, output means, and data storage means. A skilled artisan can readily appreciate that any one of the currently available computer-based systems are suitable for use in the present invention. As stated above, the computer-based systems of the present invention comprise a data storage means having stored

therein a nucleotide sequence of the present invention and the necessary hardware means and software means for supporting and implementing a search means. As used herein, "data storage means" refers to memory which can store nucleotide sequence information of the present invention, or a memory access means which can access manufactures having recorded thereon the nucleotide sequence information of the present invention.

As used herein, "search means" refers to one or more programs which are implemented on the computer-based system to compare a target sequence or target structural motif with the sequence information stored within the data storage means. Search means are used to identify fragments or regions of a known sequence which match a particular target sequence or target motif. A variety of known algorithms are disclosed publicly and a variety of commercially available software for conducting search means are and can be used in the computer-based systems of the present invention. Examples of such software includes, but is not limited to, Smith-Waterman, MacPattern (EMBL), BLASTN and BLASTA (NPOLYPEPTIDEIA). A skilled artisan can readily recognize that any one of the available algorithms or implementing software packages for conducting homology searches can be adapted for use in the present computer-based systems. As used herein, a "target sequence" can be any nucleic acid or amino acid sequence of six or more nucleotides or two or more amino acids. A skilled artisan can readily recognize that the longer a target sequence is, the less likely a target sequence will be present as a random occurrence in the database. The most preferred sequence length of a target sequence is from about 10 to 300 amino acids, more preferably from about 30 to 100 nucleotide residues. However, it is well recognized that searches for commercially important fragments, such as sequence fragments involved in gene expression and protein processing, may be of shorter length.

As used herein, "a target structural motif," or "target motif," refers to any rationally selected sequence or combination of sequences in which the sequence(s) are chosen based on a three-dimensional configuration which is formed upon the folding of the target motif. There are a variety of target motifs known in the art. Protein target motifs include, but are not limited to, enzyme active sites and signal sequences. Nucleic acid target motifs include, but are not limited to, promoter sequences, hairpin structures and inducible expression elements (protein binding sequences).

4.15 TRIPLE HELIX FORMATION

10

15

20

25

30

35

In addition, the fragments of the present invention, as broadly described, can be used to control gene expression through triple helix formation or antisense DNA or RNA, both of which methods are based on the binding of a polynucleotide sequence to DNA or RNA.

Polynucleotides suitable for use in these methods are preferably 20 to 40 bases in length and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 15241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Olmno, J. Neurochem.

56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide.

4.16 DIAGNOSTIC ASSAYS AND KITS

15

20

25

30

35

The present invention further provides methods to identify the presence or expression of one of the ORFs of the present invention, or homolog thereof, in a test sample, using a nucleic acid probe or antibodies of the present invention, optionally conjugated or otherwise associated with a suitable label.

In general, methods for detecting a polynucleotide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polynucleotide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polynucleotide of the invention is detected in the sample. Such methods can also comprise contacting a sample under stringent hybridization conditions with nucleic acid primers that anneal to a polynucleotide of the invention under such conditions, and amplifying annealed polynucleotides, so that if a polynucleotide is amplified, a polynucleotide of the invention is detected in the sample.

In general, methods for detecting a polypeptide of the invention can comprise contacting a sample with a compound that binds to and forms a complex with the polypeptide for a period sufficient to form the complex, and detecting the complex, so that if a complex is detected, a polypeptide of the invention is detected in the sample.

In detail, such methods comprise incubating a test sample with one or more of the antibodies or one or more of the nucleic acid probes of the present invention and assaying for binding of the nucleic acid probes or antibodies to components within the test sample.

Conditions for incubating a nucleic acid probe or antibody with a test sample vary.

Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the nucleic acid probe or antibody used in the assay. One skilled in the art will recognize that any one of the commonly available hybridization,

amplification or immunological assay formats can readily be adapted to employ the nucleic acid probes or antibodies of the present invention. Examples of such assays can be found in Chard, T., An Introduction to Radioimmunoassay and Related Techniques, Elsevier Science Publishers, Amsterdam, The Netherlands (1986); Bullock, G.R. et al., Techniques in Immunocytochemistry, Academic Press, Orlando, FL Vol. 1 (1982), Vol. 2 (1983), Vol. 3 (1985); Tijssen, P., Practice and Theory of immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology, Elsevier Science Publishers, Amsterdam, The Netherlands (1985). The test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as sputum, blood, serum, plasma, or urine. The test sample used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can be readily be adapted in order to obtain a sample which is compatible with the system utilized.

In another embodiment of the present invention, kits are provided which contain the necessary reagents to carry out the assays of the present invention. Specifically, the invention provides a compartment kit to receive, in close confinement, one or more containers which comprises: (a) a first container comprising one of the probes or antibodies of the present invention; and (b) one or more other containers comprising one or more of the following: wash reagents, reagents capable of detecting presence of a bound probe or antibody.

In detail, a compartment kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allows one to efficiently transfer reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated, and the agents or solutions of each container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the antibodies used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, etc.), and containers which contain the reagents used to detect the bound antibody or probe. Types of detection reagents include labeled nucleic acid probes, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. One skilled in the art will readily recognize that the disclosed probes and antibodies of the present invention can be readily incorporated into one of the established kit formats which are well known in the art.

4.17 MEDICAL IMAGING

The novel polypeptides and binding partners of the invention are useful in medical imaging of sites expressing the molecules of the invention (e.g., where the polypeptide of the invention is involved in the immune response, for imaging sites of inflammation or infection). See, e.g., Kunkel et al., U.S. Pat. NO. 5,413,778. Such methods involve chemical attachment of a labeling or imaging agent, administration of the labeled polypeptide to a subject in a pharmaceutically acceptable carrier, and imaging the labeled polypeptide in vivo at the target site.

4.18 SCREENING ASSAYS

10

15

20

25

30

35

Using the isolated proteins and polynucleotides of the invention, the present invention further provides methods of obtaining and identifying agents which bind to a polypeptide encoded by an ORF corresponding to any of the nucleotide sequences set forth in SEQ ID NO:1-1350, or bind to a specific domain of the polypeptide encoded by the nucleic acid. In detail, said method comprises the steps of:

- (a) contacting an agent with an isolated protein encoded by an ORF of the present invention, or nucleic acid of the invention; and
 - (b) determining whether the agent binds to said protein or said nucleic acid.

In general, therefore, such methods for identifying compounds that bind to a polynucleotide of the invention can comprise contacting a compound with a polynucleotide of the invention for a time sufficient to form a polynucleotide/compound complex, and detecting the complex, so that if a polynucleotide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Likewise, in general, therefore, such methods for identifying compounds that bind to a polypeptide of the invention can comprise contacting a compound with a polypeptide of the invention for a time sufficient to form a polypeptide/compound complex, and detecting the complex, so that if a polypeptide/compound complex is detected, a compound that binds to a polynucleotide of the invention is identified.

Methods for identifying compounds that bind to a polypeptide of the invention can also comprise contacting a compound with a polypeptide of the invention in a cell for a time sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a receptor gene sequence in the cell, and detecting the complex by detecting reporter gene sequence expression, so that if a polypeptide/compound complex is detected, a compound that binds a polypeptide of the invention is identified.

Compounds identified via such methods can include compounds which modulate the activity of a polypeptide of the invention (that is, increase or decrease its activity, relative to

activity observed in the absence of the compound). Alternatively, compounds identified via such methods can include compounds which modulate the expression of a polynucleotide of the invention (that is, increase or decrease expression relative to expression levels observed in the absence of the compound). Compounds, such as compounds identified via the methods of the invention, can be tested using standard assays well known to those of skill in the art for their ability to modulate activity/expression.

The agents screened in the above assay can be, but are not limited to, peptides, carbohydrates, vitamin derivatives, or other pharmaceutical agents. The agents can be selected and screened at random or rationally selected or designed using protein modeling techniques.

10

20

25

30

35

For random screening, agents such as peptides, carbohydrates, pharmaceutical agents and the like are selected at random and are assayed for their ability to bind to the protein encoded by the ORF of the present invention. Alternatively, agents may be rationally selected or designed. As used herein, an agent is said to be "rationally selected or designed" when the agent is chosen based on the configuration of the particular protein. For example, one skilled in the art can readily adapt currently available procedures to generate peptides, pharmaceutical agents and the like, capable of binding to a specific peptide sequence, in order to generate rationally designed antipeptide peptides, for example see Hurby et al., Application of Synthetic Peptides: Antisense Peptides," In Synthetic Peptides, A User's Guide, W.H. Freeman, NY (1992), pp. 289-307, and Kaspczak et al., Biochemistry 28:9230-8 (1989), or pharmaceutical agents, or the like.

In addition to the foregoing, one class of agents of the present invention, as broadly described, can be used to control gene expression through binding to one of the ORFs or EMFs of the present invention. As described above, such agents can be randomly screened or rationally designed/selected. Targeting the ORF or EMF allows a skilled artisan to design sequence specific or element specific agents, modulating the expression of either a single ORF or multiple ORFs which rely on the same EMF for expression control. One class of DNA binding agents are agents which contain base residues which hybridize or form a triple helix formation by binding to DNA or RNA. Such agents can be based on the classic phosphodiester, ribonucleic acid backbone, or can be a variety of sulfhydryl or polymeric derivatives which have base attachment capacity.

Agents suitable for use in these methods preferably contain 20 to 40 bases and are designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense - Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)). Triple helix-formation optimally results in a shut-off of RNA transcription

from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques have been demonstrated to be effective in model systems. Information contained in the sequences of the present invention is necessary for the design of an antisense or triple helix oligonucleotide and other DNA binding agents.

Agents which bind to a protein encoded by one of the ORFs of the present invention can be used as a diagnostic agent. Agents which bind to a protein encoded by one of the ORFs of the present invention can be formulated using known techniques to generate a pharmaceutical composition.

10 4.19 USE OF NUCLEIC ACIDS AS PROBES

5

15

20

25

30

Another aspect of the subject invention is to provide for polypeptide-specific nucleic acid hybridization probes capable of hybridizing with naturally occurring nucleotide sequences. The hybridization probes of the subject invention may be derived from any of the nucleotide sequences SEQ ID NO:1-1350. Because the corresponding gene is only expressed in a limited number of tissues, a hybridization probe derived from of any of the nucleotide sequences SEQ ID NO:1-1350 can be used as an indicator of the presence of RNA of cell type of such a tissue in a sample.

Any suitable hybridization technique can be employed, such as, for example, in situ hybridization. PCR as described in US Patents Nos. 4,683,195 and 4,965,188 provides additional uses for oligonucleotides based upon the nucleotide sequences. Such probes used in PCR may be of recombinant origin, may be chemically synthesized, or a mixture of both. The probe will comprise a discrete nucleotide sequence for the detection of identical sequences or a degenerate pool of possible sequences for identification of closely related genomic sequences.

Other means for producing specific hybridization probes for nucleic acids include the cloning of nucleic acid sequences into vectors for the production of mRNA probes. Such vectors are known in the art and are commercially available and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerase as T7 or SP6 RNA polymerase and the appropriate radioactively labeled nucleotides. The nucleotide sequences may be used to construct hybridization probes for mapping their respective genomic sequences. The nucleotide sequence provided herein may be mapped to a chromosome or specific regions of a chromosome using well known genetic and/or chromosomal mapping techniques. These techniques include in situ hybridization, linkage analysis against known chromosomal markers, hybridization screening with libraries or flow-sorted chromosomal preparations specific to known chromosomes, and the like. The technique of fluorescent in situ hybridization of

chromosome spreads has been described, among other places, in Verma et al (1988) Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York NY.

Fluorescent in situ hybridization of chromosomal preparations and other physical chromosome mapping techniques may be correlated with additional genetic map data. Examples of genetic map data can be found in the 1994 Genome Issue of Science (265:1981f). Correlation between the location of a nucleic acid on a physical chromosomal map and a specific disease (or predisposition to a specific disease) may help delimit the region of DNA associated with that genetic disease. The nucleotide sequences of the subject invention may be used to detect differences in gene sequences between normal, carrier or affected individuals.

4.20 PREPARATION OF SUPPORT BOUND OLIGONUCLEOTIDES

10

15

20

25

30

Oligonucleotides, *i.e.*, small nucleic acid segments, may be readily prepared by, for example, directly synthesizing the oligonucleotide by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer.

Support bound oligonucleotides may be prepared by any of the methods known to those of skill in the art using any suitable support such as glass, polystyrene or Teflon. One strategy is to precisely spot oligonucleotides synthesized by standard synthesizers. Immobilization can be achieved using passive adsorption (Inouye & Hondo, (1990) J. Clin. Microbiol. 28(6) 1469-72); using UV light (Nagata *et al.*, 1985; Dahlen *et al.*, 1987; Morrissey & Collins, (1989) Mol. Cell Probes 3(2) 189-207) or by covalent binding of base modified DNA (Keller *et al.*, 1988; 1989); all references being specifically incorporated herein.

Another strategy that may be employed is the use of the strong biotin-streptavidin interaction as a linker. For example, Broude *et al.* (1994) Proc. Natl. Acad. Sci. USA 91(8) 3072-6, describe the use of biotinylated probes, although these are duplex probes, that are immobilized on streptavidin-coated magnetic beads. Streptavidin-coated beads may be purchased from Dynal, Oslo. Of course, this same linking chemistry is applicable to coating any surface with streptavidin. Biotinylated probes may be purchased from various sources, such as, *e.g.*, Operon Technologies (Alameda, CA).

Nunc Laboratories (Naperville, IL) is also selling suitable material that could be used. Nunc Laboratories have developed a method by which DNA can be covalently bound to the microwell surface termed Covalink NH. CovaLink NH is a polystyrene surface grafted with secondary amino groups (>NH) that serve as bridge-heads for further covalent coupling. CovaLink Modules may be purchased from Nunc Laboratories. DNA molecules may be bound to CovaLink exclusively at the 5'-end by a phosphoramidate bond, allowing immobilization of more than 1 pmol of DNA (Rasmussen et al., (1991) Anal. Biochem. 198(1) 138-42).

The use of CovaLink NH strips for covalent binding of DNA molecules at the 5'-end has been described (Rasmussen et al., (1991). In this technology, a phosphoramidate bond is employed (Chu et al., (1983) Nucleic Acids Res. 11(8) 6513-29). This is beneficial as immobilization using only a single covalent bond is preferred. The phosphoramidate bond joins the DNA to the CovaLink NH secondary amino groups that are positioned at the end of spacer arms covalently grafted onto the polystyrene surface through a 2 nm long spacer arm. To link an oligonucleotide to CovaLink NH via an phosphoramidate bond, the oligonucleotide terminus must have a 5'-end phosphate group. It is, perhaps, even possible for biotin to be covalently bound to CovaLink and then streptavidin used to bind the probes.

More specifically, the linkage method includes dissolving DNA in water (7.5 ng/ul) and denaturing for 10 min. at 95°C and cooling on ice for 10 min. Ice-cold 0.1 M 1-methylimidazole, pH 7.0 (1-MeIm₇), is then added to a final concentration of 10 mM 1-MeIm₇. A ss DNA solution is then dispensed into CovaLink NH strips (75 ul/well) standing on ice.

10

15

20

25

Carbodiimide 0.2 M 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC), dissolved in 10 mM 1-MeIm₇, is made fresh and 25 ul added per well. The strips are incubated for 5 hours at 50°C. After incubation the strips are washed using, e.g., Nunc-Immuno Wash; first the wells are washed 3 times, then they are soaked with washing solution for 5 min., and finally they are washed 3 times (where in the washing solution is 0.4 N NaOH, 0.25% SDS heated to 50°C).

It is contemplated that a further suitable method for use with the present invention is that described in PCT Patent Application WO 90/03382 (Southern & Maskos), incorporated herein by reference. This method of preparing an oligonucleotide bound to a support involves attaching a nucleoside 3'-reagent through the phosphate group by a covalent phosphodiester link to aliphatic hydroxyl groups carried by the support. The oligonucleotide is then synthesized on the supported nucleoside and protecting groups removed from the synthetic oligonucleotide chain under standard conditions that do not cleave the oligonucleotide from the support. Suitable reagents include nucleoside phosphoramidite and nucleoside hydrogen phosphorate.

An on-chip strategy for the preparation of DNA probe for the preparation of DNA probe arrays may be employed. For example, addressable laser-activated photodeprotection may be employed in the chemical synthesis of oligonucleotides directly on a glass surface, as described by Fodor *et al.* (1991) Science 251(4995) 767-73, incorporated herein by reference. Probes may also be immobilized on nylon supports as described by Van Ness *et al.* (1991) Nucleic Acids Res. 19(12) 3345-50; or linked to Teflon using the method of Duncan & Cavalier (1988) Anal. Biochem. 169(1) 104-8; all references being specifically incorporated herein.

To link an oligonucleotide to a nylon support, as described by Van Ness *et al.* (1991), requires activation of the nylon surface via alkylation and selective activation of the 5'-amine of oligonucleotides with cyanuric chloride.

One particular way to prepare support bound oligonucleotides is to utilize the light-generated synthesis described by Pease *et al.*, (1994) PNAS USA 91(11) 5022-6, incorporated herein by reference). These authors used current photolithographic techniques to generate arrays of immobilized oligonucleotide probes (DNA chips). These methods, in which light is used to direct the synthesis of oligonucleotide probes in high-density, miniaturized arrays, utilize photolabile 5'-protected *N*-acyl-deoxynucleoside phosphoramidites, surface linker chemistry and versatile combinatorial synthesis strategies. A matrix of 256 spatially defined oligonucleotide probes may be generated in this manner.

4.21 PREPARATION OF NUCLEIC ACID FRAGMENTS

5

10

15

20

25

30

The nucleic acids may be obtained from any appropriate source, such as cDNAs, genomic DNA, chromosomal DNA, microdissected chromosome bands, cosmid or YAC inserts, and RNA, including mRNA without any amplification steps. For example, Sambrook *et al.* (1989) describes three protocols for the isolation of high molecular weight DNA from mammalian cells (p. 9.14-9.23).

DNA fragments may be prepared as clones in M13, plasmid or lambda vectors and/or prepared directly from genomic DNA or cDNA by PCR or other amplification methods. Samples may be prepared or dispensed in multiwell plates. About 100-1000 ng of DNA samples may be prepared in 2-500 ml of final volume.

The nucleic acids would then be fragmented by any of the methods known to those of skill in the art including, for example, using restriction enzymes as described at 9.24-9.28 of Sambrook et al. (1989), shearing by ultrasound and NaOII treatment.

Low pressure shearing is also appropriate, as described by Schriefer *et al.* (1990) Nucleic Acids Res. 18(24) 7455-6, incorporated herein by reference). In this method, DNA samples are passed through a small French pressure cell at a variety of low to intermediate pressures. A lever device allows controlled application of low to intermediate pressures to the cell. The results of these studies indicate that low-pressure shearing is a useful alternative to sonic and enzymatic DNA fragmentation methods.

One particularly suitable way for fragmenting DNA is contemplated to be that using the two base recognition endonuclease, *CviJI*, described by Fitzgerald *et al.* (1992) Nucleic Acids Res. 20(14) 3753-62. These authors described an approach for the rapid fragmentation and fractionation

of DNA into particular sizes that they contemplated to be suitable for shotgun cloning and sequencing.

The restriction endonuclease CviJI normally cleaves the recognition sequence PuGCPy between the G and C to leave blunt ends. Atypical reaction conditions, which alter the specificity of this enzyme (CviJI**), yield a quasi-random distribution of DNA fragments form the small molecule pUC19 (2688 base pairs). Fitzgerald et al. (1992) quantitatively evaluated the randomness of this fragmentation strategy, using a CviJI** digest of pUC19 that was size fractionated by a rapid gel filtration method and directly ligated, without end repair, to a lac Z minus M13 cloning vector. Sequence analysis of 76 clones showed that CviJI** restricts pyGCPy and PuGCPu, in addition to PuGCPy sites, and that new sequence data is accumulated at a rate consistent with random fragmentation.

As reported in the literature, advantages of this approach compared to sonication and agarose gel fractionation include: smaller amounts of DNA are required (0.2-0.5 ug instead of 2-5 ug); and fewer steps are involved (no preligation, end repair, chemical extraction, or agarose gel electrophoresis and elution are needed

Irrespective of the manner in which the nucleic acid fragments are obtained or prepared, it is important to denature the DNA to give single stranded pieces available for hybridization. This is achieved by incubating the DNA solution for 2-5 minutes at 80-90°C. The solution is then cooled quickly to 2°C to prevent renaturation of the DNA fragments before they are contacted with the chip. Phosphate groups must also be removed from genomic DNA by methods known in the art.

4.22 PREPARATION OF DNA ARRAYS

5

10

15

20

25

30

Arrays may be prepared by spotting DNA samples on a support such as a nylon membrane. Spotting may be performed by using arrays of metal pins (the positions of which correspond to an array of wells in a microtiter plate) to repeated by transfer of about 20 nl of a DNA solution to a nylon membrane. By offset printing, a density of dots higher than the density of the wells is achieved. One to 25 dots may be accommodated in 1 mm², depending on the type of label used. By avoiding spotting in some preselected number of rows and columns, separate subsets (subarrays) may be formed. Samples in one subarray may be the same genomic segment of DNA (or the same gene) from different individuals, or may be different, overlapped genomic clones. Each of the subarrays may represent replica spotting of the same samples. In one example, a selected gene segment may be amplified from 64 patients. For each patient, the amplified gene segment may be in one 96-well plate (all 96 wells containing the same sample). A plate for each of the 64 patients is prepared. By using a 96-pin device, all samples may be spotted on one 8 x 12 cm membrane.

Subarrays may contain 64 samples, one from each patient. Where the 96 subarrays are identical, the dot span may be 1 mm² and there may be a 1 mm space between subarrays.

Another approach is to use membranes or plates (available from NUNC, Naperville, Illinois) which may be partitioned by physical spacers e.g. a plastic grid molded over the membrane, the grid being similar to the sort of membrane applied to the bottom of multiwell plates, or hydrophobic strips. A fixed physical spacer is not preferred for imaging by exposure to flat phosphor-storage screens or x-ray films.

The present invention is illustrated in the following examples. Upon consideration of the present disclosure, one of skill in the art will appreciate that many other embodiments and variations may be made in the scope of the present invention. Accordingly, it is intended that the broader aspects of the present invention not be limited to the disclosure of the following examples. The present invention is not to be limited in scope by the exemplified embodiments which are intended as illustrations of single aspects of the invention, and compositions and methods which are functionally equivalent are within the scope of the invention. Indeed, numerous modifications and variations in the practice of the invention are expected to occur to those skilled in the art upon consideration of the present preferred embodiments. Consequently, the only limitations which should be placed upon the scope of the invention are those which appear in the appended claims.

All references cited within the body of the instant specification are hereby incorporated by reference in their entirety.

5.0 EXAMPLES

10

15

20

25

30

5.1 EXAMPLE 1

Novel Nucleic Acid Sequences Obtained From Various Libraries

A plurality of novel nucleic acids were obtained from cDNA libraries prepared from various human tissues and in some cases isolated from a genomic library derived from human chromosome using standard PCR, SBH sequence signature analysis and Sanger sequencing techniques. The inserts of the library were amplified with PCR using primers specific for the vector sequences which flank the inserts. Clones from cDNA libraries were spotted on nylon membrane filters and screened with oligonucleotide probes (e.g., 7-mers) to obtain signature sequences. The clones were clustered into groups of similar or identical sequences. Representative clones were selected for sequencing.

In some cases, the 5' sequence of the amplified inserts was then deduced using a typical Sanger sequencing protocol. PCR products were purified and subjected to fluorescent dye terminator cycle sequencing. Single pass gel sequencing was done using a 377 Applied Biosystems

(ABI) sequencer to obtain the novel nucleic acid sequences. In some cases RACE (Random Amplification of cDNA Ends) was performed to further extend the sequence in the 5' direction.

5.2 EXAMPLE 2

5 Novel Contigs

10

15

20

25

30

The novel contigs of the invention were assembled from sequences that were obtained from a cDNA library by methods described in Example 1 above, and in some cases sequences obtained from one or more public databases. The sequences for the resulting nucleic acid contigs are designated as SEQ ID NO: 1-1350 and are provided in the attached Sequence Listing. The contigs were assembled using an EST sequence as a seed. Then a recursive algorithm was used to extend the seed EST into an extended assemblage, by pulling additional sequences from different databases (i.e., Hyseq's database containing EST sequences, dbEST version 114, gb pri 114, and UniGene version 101) that belong to this assemblage. The algorithm terminated when there was no additional sequences from the above databases that would extend the assemblage. Inclusion of component sequences into the assemblage was based on a BLASTN hit to the extending assemblage with BLAST score greater than 300 and percent identity greater than 95%.

Table 3 sets forth the novel predicted polypeptides (including proteins) encoded by the novel polynucleotides (SEQ ID NO:189-282) of the present invention, and their corresponding nucleotide locations to each of SEQ ID NO: 189-282. Table 3 also indicates the method by which the polypeptide was predicted. Method A refers to a polypeptide obtained by using a software program called FASTY (available from http://fasta.bioch.virginia.edu) which selects a polypeptide based on a comparison of the translated novel polynucleotide to known polynucleotides (W.R. Pearson, Methods in Enzymology, 183:63-98 (1990), herein incorporated by reference). Method B refers to a polypeptide obtained by using a software program called GenScan for human/vertebrate sequences (available from Stanford University, Office of Technology Licensing) that predicts the polypeptide based on a probabilistic model of gene structure/compositional properties (C. Burge and S. Karlin, J. Mol. Biol., 268:78-94 (1997), incorporated herein by reference). Method C refers to a polypeptide obtained by using a Hyseq proprietary software program that translates the novel polynucleotide and its complementary strand into six possible amino acid sequences (forward and reverse frames) and chooses the polypeptide with the longest open reading frame.

The nearest neighbor results for SEQ ID NO: 1-1350 were obtained by a BLASTP version 2.0al 19MP-WashU search against Genpept release 120 and Geneseq database October 12, 2000, update 21 (Derwent), using BLAST algorithm. The nearest neighbor result showed the

closest homologue for SEQ ID NO:1-1350. The nearest neighbor results for SEQ ID NO: 1-1350 are shown in Table 2 below.

Tables 1, 2 and 3 follow. Table 1 shows the various tissue sources of SEQ ID NO: 1-1350. Table 2 shows the nearest neighbor result for the assembled contig. The nearest neighbor result shows the closest homolog with an identifiable function for each assemblage. Table 3 contains the start and stop nucleotides for the translated amino acid sequence for which each assemblage encodes. Table 3 also provides a correlation between the amino acid sequences set forth in the Sequence Listing, the nucleotide sequences set forth in the Sequence Listing and the SEQ ID NO. in USSN 09/496,914.

TABLE 1

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
adult brain	GIBCO	AB3001	111 151 188 215 662-665 877 910 927
	1 5200	1	976 1233 1319
adult brain	GIBCO	ABD003	41 49 74 101 111 120 132 141-142 151
444, 6144,	GEOC	1125003	217 225 238 271 317 404 446 469 503
	1	1	513-514 535 550 564 573 666-669 798
			898 910 927 976 1067 1083 1085 1178
		}	1254
adult brain	Clontech	ABR001	39 216 238 327 356 535 927 1056 1121
uoon oran	Cioniccii	ADROOT	1178-1180 1199 1251
adult brain	Clontech	ABR006	74 611 949 1034 1136
adult brain	Clontech	ABR008	14 32 41 61 81 86 89 120 132 138 145
dddir Drum	Cioniccii	ADROOD	147 188 197 208 225 227-239 250 300-
		1	303 312 316 328-331 340 357-362 374
	ł		380 384-391 408 414 446 448 464-467
			483 488 495-496 505 512 521 535 550
			566 571 577 585 590 594 598 634 641
			658 666 683 725 742 764 767 786 801
			805 810 823 826 829 831 836 841 887-
			923 927 934 943 950-951 963 976 995
			1000-1001 1006 1026 1034 1048 1057-
			1067 1086 1088 1090 1118 1120 1122-
	1		1128 1142 1162 1181-1192 1199 1204
	}		1218-1219 1225 1232 1253 1267 1271-
			1306 1342 1347 1349-1350
adult brain	Clontech	ABR011	49 238 1219
adult brain	BioChain	ABR012	74 238
adult brain	Invitrogen	ABR013	868 1268
adult brain	Invitrogen	ABT004	49 117 138 191 217 252 291 305 535
		l	566 596 663 670 746 798 816-819 876
) ,	892 898 922 943 963 1034-1036 1121
cultured	Strategene	ADP001	41 74 101 138 211 238 304 537 582
preadipocytes	<u>'</u>	<u> </u>	740 798 883 943 976 1067
adrenal gland	Clontech	ADR002	49 74 101 111 120 127 151 215 238
		Ĭ	240-247 316 330 363-364 404 414 534-
			535 833 924-940 950 963 976 1001
		1	1003 1067-1070 1118 1156 1193-1200
			1325
adult heart	GIBCO	AHR001	38 49 71-72 74-77 79 92 99 101 111
	}	1	118 129 132 138 151 158-163 182 195-
			203 215 217 238 264 269 353 384 398
			408 434-439 446 504 512-513 519 537
			562-573 577 611-614 616-619 658 661
		Ì	671-672 722 734 757-773 815 828-835
			874 891 898 919 926-927 976 988
			1021 1037 1041 1062 1067 1071 1080
		-	1083 1093 1122 1131 1185 1201 1254
adult kidney	GIBCO	AKD001	1308 1331 1335 41 49 51 71-74 78-85 94 100-101 103-
addit kidiley	dibeo	AKDOOL	107 111 119-120 138 151 157 215 217-
		{	218 238 250 264 294 304 384 404 440
	(446 454 477 504-505 509 514 518-519
	1		535 537 564 574-583 620-627 639 653
			673-675 705 753 789 831 844 851 859
			877 909 918 927 956 963 976 1067
			1074 1083 1095 1178 1302 1331 1335
adult kidney	Invitrogen	AKT002	11-12 41 49 111-112 215-217 294 316
		,	446 487 564 575 844 868 910 927 976
	1		1116
adult lung	GIBCO	ALG001	8 101 111 151 187 402 446 490 514
<u>-</u>		<u> </u>	

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
113000 0118			518 537 545 549 580 582 592 594 634
			640 651-652 676-678 725 851 873 918
			952 976 1042 1067 1076 1083 1152
lymph node	Clontech	ALN001	8 111 121 151 180-182 188 215 537
	ĺ		545 549 651 679-682 789 804-810 868
			873 927 952 976 1042 1059 1335
young liver	GIBCO	ALV001	8 64 79 111 186 215-216 238 446 514
			519 537 564 653 683-684 698 753 798
			813 833 840 858 927 976 1038-1039
			1051 1085 1224 1245 1256
adult liver	Invitrogen	ALV002	40 71 292-293 305 384 468-469 496
	Ì	ľ	505 657 675 714 753 832 844 941-942
		1777000	976 1040 1076 1256 1293
adult liver	Clontech	ALV003	976
adult ovary	Invitrogen	AOV001	8 32 36 38 41 49 51 71 74 79-80 101 104 111 120 122-125 138 140 143-149
			151 188-190 207-212 215-217 238 264
		1	316 384 409 440 445-446 496 504 512
	•		514 518-519 535 537 549-550 564 566
	1	1	571 580 582 600 618 638 657 667 681
			685-697 699 705 722 735-744 761 771
			815 833 842-865 868 875-876 918 926-
i			927 950 952 963 976 1023 1042 1048
			1051 1059 1072 1076 1083 1117 1120
İ			1124 1131 1144 1174 1224 1268 1331
			1335
adult placenta	Clontech	APL001	102 217 238 537 641 700
placenta	Invitrogen	APL002	663 851 1048
adult spleen	GIBCO	ASP001	8 45 74 111 132 140 151 185 217 238
			294 414 446 477 504 514 534 545 549
			592 722 873 883 952 976 1041-1042
<u> </u>		·	1083 1093-1094 1152 1224
testis	GIBCO	ATS001	72 107 111 113 126 140 151 183 215 238 446 497 537 642 701-706 811 877
			927 962 976 1083 1117 1131
adult bladder	Invitrogen	BLD001	41 151 191 402-405 409 414 496 545
addit bladder	invidogen	BLDOVI	592 607 706 873 952 1178 1329-1335
bone marrow	Clontech	BMD001	8 58-62 65-68 74 79 108 111 116 137
Jone mane.	Cidinicon	21.12001	147 151 164-174 213-215 238 305-307
	1		374 404 446 460 466 516 519 534 538-
			541 544-546 549-554 566 584 586 592
			596 607 610 628-629 643-645 652 707-
			708 774-789 844 866-871 873 919 927
			952 963 976 998 1034 1042 1064 1083
	Ì	}	1085 1120 1132 1152 1225 1229 1268
			1307 1310
bone marrow	Clontech	BMD002	6 8 37-38 52 74 77 105 111 129 132
			210 317 510-511 545 549 581 598 628
	J		638 724 766 789 844 860 868 873 919
			927 952 963 968 976 1042 1111 1141
hana	Clout-ob	BMD004	1160-1161 1229 1266 1346 111 238 282 549 1083
bone marrow adult colon	Clontech Invitrogen	CLN001	52 260 264 299 494 536 545 564 592
addit cololi	TILATOROGETI	CLIMONI	844 873 877 952 976 1042 1152 1268
	1		1336-1337
adult cervix	BioChain	CVX001	49 51 129 132 151 205 207 238 332-
	Diociani	3172001	335 365-367 392-401 440 466 470-471
			518 537 597 629 832 877 927 976 1006
	1		1085 1117 1129-1134 1192 1202-1205
			1219 1309-1328
diaphragm	BioChain	DIA002	74 976 1083

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
endothelial cells	Strategene	EDT001	32 40-41 49 74 79 101 111 120 132
			138 151 204-206 215-217 238 269 316
		ļ	414 433 505 510 513 550 555 580 582
			596 675 722 745 798 814 836-841 851
			918 976 1041 1043 1073 1083 1131
		İ	1331
Genomic clones	Genomic DNA	EPM001	525-532 927
from the short arm	from Genetic		1
of chromosome 8	Research		
Genomic clones	Genomic DNA	EPM003	47 525
from the short arm	from Genetic		!
of chromosome 8	Research		•
Genomic clones	Genomic DNA	EPM004	525 927
from the short arm	from Genetic		
of chromosome 8	Research		
Genomic clones	Genomic DNA	EPM005	531
from the short arm	from Genetic		ĺ
of chromosome 8	Research		
esophagus	BioChain	ESO002	74 138 238
fetal brain	Clontech	FBR001	441-442 927
fetal brain	Clontech	FBR004	215 893 927 1001
fetal brain	Clontech	FBR006	48 61 101 120 132 138 140 147 208
	j		225 271 317 319 336 359 368 405-414
			519 550 571 594 686 715 722 764 824
	1	j	829 836 859 909 927 943 947 963 1057
	İ		1067-1068 1104 1135-1140 1162 1206-
	}		1207 1235 1268 1288 1307-1308 1319
	1		1338-1350
fetal brain	Clontech	FBRs03	111 446
fetal brain	Invitrogen	FBT002	41 51 120 151 192-194 264 504 512
			535 683 761 798 820-827 844 876 909
		·	963 976 1026 1048 1083 1144 1302
fetal heart	Invitrogen	FHR001	446 566 761
fetal kidney	Clontech	FKD001	51 74 111 127 140 151 184 294 537
			550 630-631 1319
fetal kidney	Clontech	FKD002	111 976 1083
fetal kidney	Invitrogen	FKD007	238 974
fetal lung	Clontech	FLG001	463 566 976 1074 1083 1093
fetal lung	Invitrogen	FLG003	41 238 330 407 415-416 537 573 844
			859 1048 1083 1116 1192
fetal liver-spleen	Columbia	FLS001	8 14 34-35 37 41 43 49 51 54-56 63-64
(University		69-71 74 77 79 87-90 101 107 110-111
		Ì	114 120 128-131 138 140 147 150-155
Ì	ì		197 210 215 217 225 238 312 367 384
			414 440 446 460 468 483 496 504-507
ļ	1		511-515 518-519 523 533-535 537 541
ł	1	i	544-545 547-550 555-560 564 566 571
			577 582 585-586 598 636 646-647 649
1	-		652 664 698 709-710 714 722-723 731
}	j] .	735-736 746-753 761 784 798 823 829
	1		832 844 851 858-859 868 873 876 898
	1	1	927 943 949 952 963 976 984 1002
			1021 1023 1040 1042 1044 1050 1083
	1		1093 1116 1120 1129 1131 1144 1174
1	}	1	1217 1251 1254 1256 1302 1308 1311
Catal Hanne and an	Columbia	ET 5000	1319
fetal liver-spleen	Columbia	FLS002	8 36-37 41-46 49 54 64 71 74 79 101
	University		111 120 129 147 207 210 215-216 238
1		 	250 330 353 359 366 383-384 414 478 505 508-509 511 515-524 534-535 537
ļ			544-545 564 566 571 577 591 598 638
L	<u></u>		מכם מכנ וכל וול וול מטל המל להל-בהל ו

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
		1,0004 2,001 1,12,10	663 671 698 714 722 725 727 751 798
			851 859 873 876 909 927 949 952 983-
			984 1002 1023 1042-1044 1085 1095
	1		1131 1144 1178 1199 1233 1240-1270
		<u> </u>	1331 1340
fetal liver-spleen	Columbia University	FLS003	64 535 976 1256
fetal liver	Invitrogen	FLV001	8 101 120 138 217 446 468 535 566
			580 722 730 749 844 918 943 976 1051
		<u> </u>	1256 1331
fetal liver	Clontech	FLV004	537 926 1256
fetal muscle	Invitrogen	FMS001	51 111 264 312 369-370 404 417-421
			425 535 537 577 598 614 836 857 1141 1208 1268
fetal muscle	Invitrogen	FMS002	537
fetal skin	Invitrogen	FSK001	13-26 32 41 51 89 107 111 147 151
			225 264 316 405 422-429 488-494 496
]		519 534-535 537 566 675 732 859 876-
	İ		877 898 947 949-950 963 976 1001
			1062 1076 1083 1117 1144 1165 1268
			1281
fetal skin	Invitrogen	FSK002	537 812
fetal spleen	BioChain	FSP001	87 549
umbilical cord	BioChain	FUC001	27-33 41 49 151 215 238 248-249 301
			316 446 495-503 519 521 534-535 537
			582 634 691 877 883 927 944-950 963
			976 1001 1075 1142-1143 1171 1218 1243 1308
fetal brain	GIBCO	HFB001	41 49 57 79 87 103 111 120 132-135
iciai oram	GIECO	111 15001	138 145 151 188 197 207 215 238 264
		,	271 294 316 367 414 440 446 466 504
].	513-514 535 542-543 550 564 571 596
		}	635 648-654 675 711-715 722-723 798
	•		832 872 876 883 927 976 1095 1144
	<u> </u>		1168 1171 1178 1211 1335
macrophage	Invitrogen	HMP001	238
infant brain	Columbia	IB2002	49-50 77 81 89 105 111 136-138 140
	University		151 161 175-179 185 216-217 264 295
		1	299 308-310 371-373 462 476 504 511- 513 533 537 564 566 571 655-657 662
			683 716-720 723 752 790-803 829 832
		•	858-859 876 898 909 949 976 1045-
			1047 1076-1087 1090 1093 1116 1122
			1144 1209-1213 1225 1233 1256 1319
			1341
infant brain	Columbia	IB2003	41 50 77 104 132 215 238 508 512-513
	University	ļ	519 566 655 714 794 918 943 976 1067
	<u> </u>	<u> </u>	1092-1093 1233
infant bra <u>in</u>	Columbia	IBM002	311 472-473 753 1214
infant brain	University	TD0001	51 111 000 100 000 000 100 1111
unant otalii	Columbia University	IBS001	51 111 376 474 790 876 949 1144 1204 1 1221
lung, fibroblast	Strategene	LFB001	151 316 462 514 534 582 675 939 1131
lung tumor	Invitrogen	LGT002	1-7 41 74 79 94 115 120 138-139 156
P emme A1	111111111111111111111111111111111111111	DOTOUL	215 217 269 280 296 337 374-375 384
	}		404 446 454 475-480 498 514 518-519
	1		522 537 545 564 577 597 653 658 705
	1		721-724 754-756 779 859 868 872-874
	- [876-877 919 927 949 951-952 959 976
			876-877 919 927 949 951-952 959 976 1002 1042 1048-1053 1076 1083 1088- 1089 1131 1144-1147 1216-1218 1229

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			1293 1311
lymphocytes	ATCC	LPC001	41 74 111 132 151 253 316 446 550
1	ļ		634 844 927 976 1085 1268
leukocyte	GIBCO	LUC001	8 11 41 74 86 91-98 101 109 111 120
		1	147 151 212 215 218 238 252 288 312-
			314.316 338 359 408 427 443-447 505
			510 512 514 518 534 545 549-550 561
	}	1	564 566 571 577 580 582 587-609 615
			632-638 658-659 698 714 725-728 832
			836 841 859 866 873-874 882-883 918-
			919 927 943 952 963 976 1042 1076
			1083 1090 1148 1152 1168 1195 1219-
	ļ		1220 1224
leukocyte	Clontech	LUC003	74 100 215 232 238 339-341 446 545
	! !		657 660 729 873 883 927 952 963 1008
			1042 1116 1120 1149-1150 1215 1222
Melanoma from cell	Clontech	MEL004	210 215 238 342 534 545 592 722 873
line ATCC #CRL	[919 929 939 952 976 1071 1118 1218
1424			1235 1245
mammary gland	Invitrogen	MMG001	8-10 40-41 49 73 80 114 138-140 147
	[217 250-256 264 297-299 305 377-378
			398 446 481-486 505 512 537 545 549
			571 592 725 730-733 816 829 836 844
			868 873 876-877 898 926 943 951-960
			963 976 995 1034 1042 1048 1054-
			1055 1076 1083 1091 1093 1116-1117
		<u> </u>	1124 1152 1302
induced neuron cells	Strategene	NTD001	39 101 111 138 238 361 1225 1251
retinoid acid induced	Strategene	NTR001	1319 74 225 976
neuronal cells	Strategene	Nikooi	14 223 910
neuronal cells	Strategene	NTU001	129 225 238 304 313 361 657 976
pituitary gland	Clontech	PIT004	976
placenta	Clontech		38 976
prostate	Clontech	PRT001	111 188 238 257-258 564 724 961-966
product			1067 1095
rectum	Invitrogen	REC001	238 430-431 841 859 868 963 1001
		1	1116
salivary gland	Clontech	SAL001	8 151 402 432-433 446 496 868 952
			976 1083 1120 1151 1184
small intestine	Clontech	SIN001	8 101 147 215 259-266 446 462 505
			545 592 660 789 836 866 873 927 952
			963 967-978 1042 1120 1152 1223-
			1224
skeletal muscle	Clontech	SKM001	238 302 927 943 992 1031
spinal cord	Clontech	SPC001	74 111 132 151 215-216 238 264 267-
			270 343-344 353 379 516 537 566 740
]			828 927 976 979-994 1092 1153-1159
			1225 1250
adult spleen	Clontech	SPLc01	698 859 1042
stomach	Clontech	STO001	210 238 271-272 537 580 705 918 952
	<u> </u>		995 1171
thalamus	Clontech	THA002	61 219-220 273-276 312 315 330 596
		<u> </u>	963 996-1007 1059 1093 1160-1162
thymus	Clonetech	THM001	8 120 151 208 221 316-317 353 639
			750 867 874 878-881 927 963 1023
			1083 1094-1096 1124
thymus	Clontech	THMc02	8 61 114 129 132 210 225 231 306
			317-319 336 340 359 380 398 446 448-
1		1	463 512 519 545 554 587 598 698 724-
			725 789 812 836 868 873 927 947 952

Tissue Origin	RNA Source	Hyseq Library Name	SEQ ID NOS:
			976 1007 1042 1083 1085 1097-1116
1]		1122 1147 1177 1226-1229 1234 1311
	1		1313
thyroid gland	Clontech	THR001	14 41 49 76 94 111 144 151 183 188
		ļ	210 217 222 253 264 271 277-286 294
	1		320-326 345-352 361 381-382 446 467
1	1		483 514 534 549-550 564 578 602 649
			844 882-883 927 950 956 976 1008-
	1		1028 1076 1083 1117-1120 1142 1163-
			1175 1230-1238 1308
trachea	Clontech	TRC001	223-225 238 287 353-354 514
ļ)	1	545 592 611 873 883-884 927
			952 1029-1031 1042 1151-1152
}			1170 1176-1177 1239
uterus	Clontech	UTR001	151 226 288-290 355 537 877
1	}		885-886 976 1001 1032-1033
			1232

TABLE 2

SEQ	Accession	Species	Description	Smith-	%
ID NO:	No.	1		Waterman Score	Identity
1	B02829	Homo sapiens	Human G protein coupled receptor hRUP5 protein SEQ ID NO:10.	460	100
2	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	111	51
3	R26173	Homo sapiens	Part of Major Yo paraneoplastic antigen (CDR62) encoded by clone pY2.	293	76
4	L29536	Homo sapiens	calcium channel L-type alpha 1 subunit	191	65
5	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	251	50
6	M11507	Homo sapiens	transferrin receptor	120	95
7	AF099100	Homo sapiens	WD-repeat protein 6	1941	93
8	Y92338	Homo sapiens	Human cancer associated antigen precursor from clone NY-REN-45.	245	82
9	G01343	Homo sapiens	Human secreted protein, SEQ ID NO: 5424.	226	91
10	AJ133798	Homo sapiens	copine VII protein	1127	68
11	G02449	Homo sapiens	Human secreted protein, SEQ ID NO: 6530.	584	99
12	X98330	Homo sapiens	ryanodine receptor 2	282	78
13	AL024498	Homo sapiens	dJ417M14.2 (novel serine/threonine-protein kinase (ortholog of mouse and rat MAK (male germ cell-associated kinase))	293	100
14	AF045577	Pan troglodytes	olfactory receptor OR93Ch	191	36
15	G03131	Homo sapiens	Human secreted protein, SEQ ID NO: 7212.	93	39
16	U26595	Rattus norvegicus	prostaglandin F2a receptor regulatory protein precursor	569	89
17	B08918	Homo sapiens	Human secreted protein sequence encoded by gene 28 SEQ ID NO:75.	99	44
18	Y36203	Homo sapiens	Human secreted protein #75.	165	75
19	U15647	Mus musculus	reverse transcriptase	106	40
20	G02701	Homo sapiens	Human secreted protein, SEQ ID NO: 6782.	544	100
21	Y35923	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 172.	1691	100
22	G04030	Homo sapiens	Human secreted protein, SEQ ID NO: 8111.	380	96
23	G02455	Homo sapiens	Human secreted protein, SEQ ID NO: 6536.	123	50
24	AF036329	Homo sapiens	gonadotropin-releasing hormone precursor, second form	284	90
25	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	96	32
26	S80119	Rattus sp.	reverse transcriptase homolog	100	34
27	U83303	Homo sapiens	line-1 reverse transcriptase	101	35
28	G03267	Homo sapiens	Human secreted protein, SEQ ID NO: 7348.	135	45

SEQ	Accession	Species	Description	Smith-	%
ID	No.	1)	Waterman	Identity
NO:		 		Score	<u> </u>
29	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	83	42
30	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	116	72
31	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	96	67
32	G03224	Homo sapiens	Human secreted protein, SEQ ID NO: 7305.	58	32
33	Y66688	Homo sapiens	Membrane-bound protein PRO1152.	2457	98
34	Y87071	Homo sapiens	Human secreted protein sequence SEQ ID NO:110.	348	95
35	U15131	Homo sapiens	p126	182	48
36	Y73464	Homo sapiens	Human secreted protein clone yl4_1 protein sequence SEQ ID NO:150.	982	90
37	AL133215	Homo sapiens	bA108L7.6 (semaphorin 4G (sema domain, immunoglobulin domain (Ig), transmembrane domain (TM) and short cytoplasmic domain))	687	99
38	AC067969	amino acids 3338-4088	Homo sapiens ryanodine receptor 1 (skeletal)	386	66
39	ÁL031588	Homo sapiens	dJ1163J1.1 (mostly supported by GENSCAN, FGENES and GENEWISE)	493	76
40	G03628	Homo sapiens	Human secreted protein, SEQ ID NO: 7709.	110	51
41	AF132969	Homo sapiens	CGI-35 protein	228	68
42	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	220	88
43	X61048	Hydra sp.	mini-collagen	105	35
44	M76546	Helianthus annuus	hydroxyproline-rich protein	110	31
45	U82288	Caenorhabditi s elegans	Rac-like GTPase	139	70
46	G03477	Homo sapiens	Human secreted protein, SEQ ID NO: 7558.	118	58
47	AF090942	Homo sapiens	PRO0657	113	63
48	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	90	59
49	AJ005560	Mus musculus	SPR2B protein	72	56
50	G02450	Homo sapiens	Human secreted protein, SEQ ID NO: 6531.	385	98
51	Y91649	Homo sapiens	Human secreted protein sequence encoded by gene 60 SEQ ID NO:322.	973	94
52	U93563	Homo sapiens	putative p150	105	38
53	Y55927	Homo sapiens	Human STLK2 protein.	699	85
54	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	145	56
55	AB008175	Mus musculus	hepatic nuclear factor 1-beta short form	356	74
56	M68941	Homo sapiens	protein-tyrosine phophatase	165	41
57	AL031600	Homo sapiens	c390E6.1 (chloride channel 7)	338	76
58	AF011417	Mus musculus	putative pheromone receptor	143	55
59	AF167320	Mus musculus	zinc finger protein ZFP113	558	68
60	U73036	Homo sapiens	interferon regultory factor 7	263	96
61	X07984	Mus musculus	protein-tyrosine kinase	297	69
62	Y29861	Homo sapiens	Human secreted protein clone cb98_4.	791	98
63	U35376	Homo sapiens	repressor transcriptional factor	485	65
64	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	785	74
65	G03883	Homo sapiens	Human secreted protein, SEQ ID NO: 7964.	88	95
66	AF177390	Manduca sexta	antennal specific membrane protein AMP	274	54
67	AB040800	Homo sapiens	SREB2	614	100
68	AF030027	Equine herpesvirus 4	24	213	26
69	G02965	Homo sapiens	Human secreted protein, SEQ ID NO: 7046.	261	95
70	W75770	Homo sapiens	Human oxidoreductase YTFO3.	1144	98
71	AB011135	Homo sapiens	KIAA0563 protein	239	76
72	AB014885	Halocynthia roretzi	HrPOPK-1	813	78
73	AF045454	Cavia · porcellus	phospholipase B	955	73
74	J02870	Mus	laminin receptor	308	61

SEQ	Accession	Species	Description	Smith-	1 %
D D	No.	Species	2-551.p.101.	Waterman	Identity
NO:	}			Score	
	 	musculus			
75	Y00826	Rattus	gp210 (AA 1-1886)	413	84
		norvegicus	[Si	Ì	
76	AF117754	Homo sapiens	thyroid hormone receptor-associated protein	351	54
	1		complex component TRAP240		1
77	Y38422	Homo sapiens	Human secreted protein.	468	76
78	Y14596	Homo sapiens	Human T-type voltage-gated Ca channel alpha-	1357	99
		<u> </u>	1-I (hCavT3).	<u></u>	1
79	Y14591	Human	APM-1 protein	767	100
		papillomaviru		l	
		s type 68		<u> </u>	
80	AL137802	Homo sapiens	dJ798A10.2 (KIAA0445 protein)	71	34
81	AP000383	Arabidopsis	protein arginine N-methyltransferase-like protein	359	65
-00	1.45015	thaliana	Data Line in the control of the cont	005	76
82	L46815	Mus	DNA binding protein Rc	895	75
02	C01600	musculus	Human asserted restain SEO ID NO. 5691	315	96
83 84	G01600 Y53886	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 5681. A suppressor of cytokine signalling protein	538	71
04	1 33000	nomo sapiens	designated HSCOP-6.	220	''
85	AB029002	Homo sapiens	KIAA1079 protein	134	42
86	Y28678	Homo sapiens	Human cw272 7 secreted protein.	325	62
87	Y99368	Homo sapiens	Human PRO1326 (UNQ686) amino acid	156	48
٥,	1	110mic suprems	sequence SEQ ID NO:100.	1	1 "
88	AJ225124	Mus	hyperpolarization-activated cation channel,	487	95
		musculus	HAC3		1
89	AF177203	Homo sapiens	cerebral cell adhesion molecule	290	56
90	Y28280	Homo sapiens	Human G-protein coupled receptor GRIR-2.	326	79
91	L39891	Homo sapiens	polycystic kidney disease-associated protein	1751	95
92	AF064876	Homo sapiens	ion channel BCNG-1	953	99
93	AF170723	Homo sapiens	protein kinase STK10	401	53
94 .	X13292	Trypanosoma	GPI-phospholipase C (AA 1 - 358)	151	37
	J	brucei			
95	Y34127	Homo sapiens	Human potassium channel K+Hnov11.	661	99
96	X03638	Rattus	sodium channel protein I (aa 1-2009)	1775	92
^=	1	norvegicus		1005	100
97	AF134213	Homo sapiens	ubiquitin-specific protease	1995 213	99
98 99	G00838 AF021935	Homo sapiens Rattus	Human secreted protein, SEQ ID NO: 4919. mytonic dystrophy kinase-related Cdc42-binding	675	48
33	AF021933	norvegicus	kinase	0/3	40
100	AF279265	Homo sapiens	putative anion transporter 1	867	98
101	AC007878	Homo sapiens	match to nuclear protein, NP220; note: sequence	160	160
101	10007070	Tromo suprems	difference at residue 58	1.00	1 00
102	U22829	Mus	P2Y purinoceptor	264	42
		musculus			
103	Y45023	Homo sapiens	Human sensory transduction G-protein coupled	516	99
		<u></u>	receptor-B3.		
104	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	787	98
105	Y87342	Homo sapiens	Human signal peptide containing protein HSPP-	343	57
	.1	<u> </u>	119 SEQ ID NO:119.		
106	AF169312	Homo sapiens	hepatic angiopoietin-related protein	212	67
107	AF116657	Homo sapiens	PRO1310	74	52
108	AE000401	Escherichia	sialic acid transporter	587	96
• • • •	1,,,,,,,	∞li		(00	100
109	Y38395	Homo sapiens	Human secreted protein encoded by gene No. 10.	693	100
110	Y78801	Homo sapiens	Hydrophobic domain containing protein clone	182	94
111	725525	Uo	HP00631 amino acid sequence.	161	95
111	Z25535	Homo sapiens	nuclear pore complex protein hnup153	274	85
112	Y94939	Homo sapiens	Human secreted protein clone ye90_1 protein sequence SEQ ID NO:84.	4/4	31
113	A F()16365	Homo sapiens	hexokinase 1 isoform td	301	71
114	AF016365 AC007956	Homo sapiens	unknown	520	75
114	M83738	Homo sapiens	protein-tyrosine phosphatase	251	92
116	AL157952	Homo sapiens	dJ875K15.1.1 (ets homologous factor (ets-	484	91
110	ALL 1772	rionio sapiens	domain transcription factor ESE-3A, isoform 1))	707	1"
117	W18084	Homo sapiens	Human Aurora-2.	546	. 87
_ <u></u> _		1		<u> </u>	

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:		ļ		Score	
118	L41816	Homo sapiens	cam kinase I	407	62
119	AJ006710	Rattus norvegicus	phosphatidylinositol 3-kinase	627	93
120	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor, PDPr	1646	94
121	S39392	Homo sapiens	protein tyrosine phosphatase, PTPase {EC 3.1.3.48}	373	68
122	U60805	Homo sapiens	oncostatin-M specific receptor beta subunit	262	88
123	Y44403	Homo sapiens	Human truncated tankyrase-1.	111	35
124	U88167	Caenorhabditi s elegans	contains similarity to C2 domains	219	29
125	AF300648	Homo sapiens	guanine nucleotide binding protein beta subunit 4	693	90
126	AB021861	Mus musculus	apoptosis signal-regulating kinase 2	153	65
127	AF305210	Homo sapiens	concentrative Na+-nucleoside cotransporter hCNT3	807	97
128	M90360	Homo sapiens	protein kinase	220	73
129	D32202	Homo sapiens	alpha 1C adrenergic receptor isoform 2	574	86
130	AF208043	Homo sapiens	IFI16b	496	67
131	AF201734	Mus musculus	testis specific serine kinase-3	800	87
132	AF112886	Bos taurus	differentiation enhancing factor 1	159	74
133	AJ278314	Homo sapiens	phospholipase C-beta-1b	554	85
134	W74802	Homo sapiens	Human secreted protein encoded by gene 73 clone HSQEL25.	1157	87
135	AB020335	Homo sapiens	Pancreas-specific gene	668	96
136	W80408	Homo sapiens	A secreted protein encoded by clone dt674_2.	866	98
137	AC002563	Homo sapiens	putative RHO/RAC effector protein; 95% similarity to P49205 (PID:g1345860)	5041	99
138	Y96736	Homo sapiens	PRO3434, a novel secreted protein.	891	100
139	AB024034	Arabidopsis thaliana	DNA-damage inducible protein DDII-like	147	55
140	W97809	Homo sapiens	Human GTPase regulator GRAF.	248	56
141	Y51557	Homo sapiens	Human PLA2 protein.	125	46
142	AF090113	Rattus norvegicus	AMPA receptor binding protein	623	93
143	W26642	Homo sapiens	Human RECK cancer-inhibiting protein.	641	82
144	U87306	Rattus norvegicus	transmembrane receptor UNC5H2	578	84
145	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	727	92
146	W63683	Homo sapiens	Human secreted protein 3.	140	40
147	M96264	Homo sapiens	galactose-1-phosphate uridyl transferase	513	81
148	D64014	Escherichia coli	HrsA	818	90
149	M83316	Escherichia coli	pppGpp phosphohydrolase	915	95
150	AL163279	Homo sapiens	homolog to cAMP response element binding and beta transducin family proteins	1261	99
151	AF179867	Homo sapiens	STE20-like kinase	940	99
152	R95332	Homo sapiens	Tumor necrosis factor receptor 1 death domain ligand (clone 3TW).	392	61
153	AF151859	Homo sapiens	CGI-101 protein	370	92
154	X66957	Homo sapiens	hexokinase type 1	489	81
155	Y16355	Homo sapiens	alternatively spliced form	432	92
156	G00857	Homo sapiens	Human secreted protein, SEQ ID NO: 4938.	349	78
157	AF159455	Mus musculus	zinc finger protein	352	74
158	L76191	Homo sapiens	interleukin-1 receptor-associated kinase	537	76
159	AP001743	Homo sapiens	putative gene, ankirin like, possible dual specifity Ser/Thr/Tyr kinase domain	670	98
160	AJ250425	Rattus norvegicus	Collybistin I	556	74
161	G02885	Homo sapiens	Human secreted protein, SEQ ID NO: 6966.	370	100

SEO	Accession	Species	Description	Smith-	T%
ID ID	No.	Operics	Description	Waterman	Identity
NO:	140.	ţ		Score	100000
162	Z22968	Homo sapiens	M130 antigen	610	100
163	AF181121	Homo sapiens	ATP-dependent Ca2+ pump PMR1	336	92
	AF055636	Homo sapiens	leucine-rich glioma-inactivated protein precursor	455	94
164	AF160798	Rattus		700	96
165		norvegicus	calcium transporter CaT1		
166	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	327	45
167	Y48607	Homo sapiens	Human breast tumour-associated protein 68.	1072	99
168	AB020741	Mus musculus	NIK-related kinase	197	43
169	AF252293	Homo sapiens	PAR3	596	44
170	U59429	Cricetinac gen. sp.	diacylglycerol kinasc eta	481	82
171	AF035268	Homo sapiens	phosphatidylserine-specific phospholipase A1	386	42
172	AF127085	Mus	semaphorin cytoplasmic domain-associated	507	82
173	Y27918	Homo sapiens	protein 3B Human secreted protein encoded by gene No.	653	99
	†		123.		<u> </u>
174	G02979	Homo sapiens	Human secreted protein, SEQ ID NO: 7060.	538	97
175	U36488	Mus musculus	embryonic stem cell phosphatase	168	55
176	W95629	Homo sapiens	Homo sapiens secreted protein gene clone gm196_4.	1022	100
177	AF289023	Homo sapiens	formiminotransferase cyclodeaminase form D	255	93
178	X04936	Homo sapiens	T-cell receptor alpha-chain (413 is 2nd base in	710	199
		<u> </u>	codon)	175	1 80
179	AF127481	Homo sapiens	non-ocogenic Rho GTPase-specific GTP exchange factor]
180	G00978	Homo sapiens	Human secreted protein, SEQ ID NO: 5059.	517	94
181	Y66645	Homo sapiens.	Membrane-bound protein PRO1310.	671	96
182	AF110640	Homo sapiens	orphan seven-transmembrane receptor	862	100
183	AB020854	Bos taurus	orphan transporter short splicing variant	766	84
184	AF169691	Homo sapiens	cadherin-like protein VR8	375	38
185	AF126372	Homo sapiens	thyrotropin-releasing hormone degrading ectoenzyme	985	99
186	L20966	Homo sapiens	phosphodiesterase	541	76
187	G02920	Homo sapiens	Human secreted protein, SEQ ID NO: 7001.	254	93
188	Y94918	Homo sapiens	Human secreted protein clone dd504_18 protein sequence SEQ ID NO:42.	301	98
189	Y66713	Homo sapiens	Membrane-bound protein PRO1309.	694	100
190	G03244	Homo sapiens	Human secreted protein, SEQ ID NO: 7325.	331	73
191	U36771	Rattus norvegicus	sn-glycerol 3-phosphate acyltransferase	707	92
192	R05935	Homo sapiens	Secreted GPIIb subunit of multiple subunit polypeptide (MSP)GPIIb-IIIa.	157	72
193	M92084	Theileria parva	casein kinase II alpha subunit	364	50
194	Y66645	Homo sapiens	Membrane-bound protein PRO1310.	448	1 90
195	W95631	Homo sapiens	Homo sapiens secreted protein gene clone hi968 2.	382	49
196	AF255614	Rattus	scaffolding protein SLIPR	680	99
197	AC021640	norvegicus Arabidopsis	putative phosphatidate phosphohydrolase	300	41
100	1 1000000	thaliana		100	12
198	AF073967	Mus musculus	olfactory receptor	316	43
L	1770-7-0	domesticus		1	100
199	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	617	98
200	AF117948	Homo sapiens	pancreas-enriched phospholipase C	625	89
201	AF128625	Homo sapiens	CDC42-binding protein kinase beta	636	94
202	AF117946	Homo sapiens	Link guanine nucleotide exchange factor II	1303	100
203	Y53021	Homo sapiens	Human secreted protein clone qc646_1 protein sequence SEQ ID NO:48.	701	99
204	AF227968	Homo sapiens	SH2-B beta signaling protein	182	79
205	S81752	Homo sapiens	DPH2L=candidate tumor suppressor gene	375	100

SEQ	Accession	Species	Description	Smith-	1%
ID ID	No.) opens	2001paon	Waterman	Identity
NO:	110.			Score	
140.	 	 	{ovarian cancer critical region of deletion}	-	
206	U18315	Sus scrofa	parathyroid receptor	122	60
207	AF255342	Homo sapiens	putative pheromone receptor V1RL1 long form	170	96
208	S52051	Rattus sp.	neurotransmitter transporter	715	94
209	W63683	Homo sapiens	Human secreted protein 3.	840	99
			similar to Drosophila photoreceptor cell-specific	541	82
210	D79992	Homo sapiens		341	02
	17115010	 	protein, calphotin.	1348	99
211	AF117948	Homo sapiens	pancreas-enriched phospholipase C		69
212	U81035	Rattus	ankyrin binding cell adhesion molecule	471	99
	1	norvegicus	neurofascin	700	1
213	AF154846	Homo sapiens	zinc finger protein	798	56
214	AF102777	Mus	FYVE finger-containing phosphoinositide kinase	933	93
	ļ	musculus			
215	AL163303	Homo sapiens	putative gene containing transmembrane domain	523	89
216	U26595	Rattus	prostaglandin F2a receptor regulatory protein	563	78
	1	norvegicus	precursor		<u> </u>
217	G04095	Homo sapiens	Human secreted protein, SEQ ID NO: 8176.	644	98
218	X75756	Homo sapiens	protein kinase C mu	314	81
219	Y66723	Homo sapiens	Membrane-bound protein PRO1100.	770	98
220	D88577	Mus	Kupffer cell receptor	567	40
	}	musculus	•	ì	1
221	AF258465	Homo sapiens	OTRPC4	853	100
222	AF021935	Rattus	mytonic dystrophy kinase-related Cdc42-binding	636	96
		norvegicus	kinase		1
223	AL136527	Homo sapiens	bA215B13.1 (A kinase (PRKA) anchor protein	693	100
			11)		
224	AB032417	Homo sapiens	WNT receptor Frizzled-4	690	99
225	AF030430	Mus	semaphorin VIa	703	68
223	1.11 050 150	musculus	ourspirot	1	
226	AE000218	Escherichia	putative dihydroxyacetone kinase (EC 2.7.1.2)	297	39
220	ALOUZIG	coli	puntive uniyaroxyaccione kumbe (De 2.7.1.2)	1 /	
227	AF302150	Homo sapiens	phosphoinositol 3-phosphate-binding protein-2	2080	100
228	AB024573	Mus	GTP-binding like protein 2	265	88
220	AB024373	musculus	O11 -binding fixe protein 2	203	1 00
229	AF122924	Xenopus	, Wnt inhibitory factor-1	316	40
223	AF 122524	laevis	, with highlottory factor-1	3.0	100
230	G03205	Homo sapiens	Human secreted protein, SEQ ID NO: 7286.	229	100
231	X98260	Homo sapiens	M-phase phosphoprotein 11	265	92
232	R92754		Human growth differentiation factor-12.	682	95
		Homo sapiens		290	100
233	R75111	Homo sapiens	Glycosyl-phosphatidylinositol-specific	290	100
	111111111	 	phospholipase-D.	025	
234	W69431	Homo sapiens	Human secreted protein cw1233_3.	235	97
235	Y08686	Homo sapiens	serine palmitoyltransferase, subunit II	859	81
236	AF118275	Homo sapiens	atrophin-related protein ARP	117	37
237	X81466	Mus	Embryo Brain Kinase	460	62
		musculus		<u> </u>	
238	U64857	Caenorhabditi	similar to the BPTI/Kunitz family of inhibitors;	284	33
		s elegans	most similar to tissue factor pathway inhibitor		1
	ļ	ļ	precursor (TFPI)	1	1
239	AJ250840	Mus	serine/threonine protein kinase	739	63
	1	musculus		1	
240	AJ223472	Mus	transcription elongation factor TFIIS.h	222	38
	1	musculus	,	ĺ	ſ
241	Y94906	Homo sapiens	Human secreted protein clone rb649 3 protein	353	52
•			sequence SEQ ID NO:18.	1	
242	AF169301	Homo sapiens	Na+/sulfate cotransporter SUT-1	591	99
243	L22022	Rattus	orphan transporter v7-3	667	93
_+3		norvegicus	a-b	1	
244	AF016191	Rattus	potassium channel	1043	98
£77	LT 010131	norvegicus	homotem commer		1 ~
245	A E007266		cone codium calcium notoscium auchanosci	645	98
245	AF097366	Homo sapiens	cone sodium-calcium potassium exchanger		98
246	Y29868	Homo sapiens	Human secreted protein clone pp325_9.	497	
247	AF180475	Homo sapiens	Not4-Np	188	83
	Y17227	Homo sapiens	Human secreted protein (clone ya1-1).	690	99
248 249	AF250910	Manduca	death-associated small cytoplasmic leucine-rich	182	31

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	1	[- · · · · · · · · · · · · · · · · · ·	Waterman	Identity
NO:		1		Score	
	 	sexta	protein SCLP		
250	AF192756	Kaposi's	Orf73	134	34
		sarcoma-			1
		associated			
	Į	herpesvirus		· L	1
251	AB022694	Homo sapiens	MOK protein kinase	209	83
252	W55045	Homo sapiens	Neural adhesion molecule (cthb0018f2 product).	469	100
253	1.46815	Mus	DNA binding protein Rc	251	67
233	D40013	musculus	DIVA binding protein Re	231	07
254	W68505	Homo sapiens	Human acid sensing ionic channel.	173	82
255	AF070066		Citron-K kinase	1201	
233	AF070000	Mus	Citron-K kinase	1201	98
256	G02491	musculus	11	460	100
256		Homo sapiens	Human secreted protein, SEQ ID NO: 6572.		100
257	Z12841	Oryctolagus	Phospholipase	368	80
000	1/05/06	cuniculus	1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	1.055	ļ
258	Y95436	Homo sapiens	Human calcium channel SOC-3/CRAC-2.	1857	99
259	AJ222968	Mus	L-periaxin	430	72
		musculus		L	
260	AJ250839	Homo sapiens	serine/threonine protein kinase	861	100
261	AJ249977	Homo sapiens	AMP-activated protein kinase gamma 3 subunit	758	98
262	AF141386	Rattus	SLIT-2	198	40
		norvegicus		.1	1
263	AF022859	Homo sapiens	neuropilin-2(a0)	335	62
264	AF160477	Homo sapiens	Ig superfamily receptor LNIR precursor	387	91
265	Y44662	Homo sapiens	Human 14273 G-protein coupled receptor	636	99
			(GPCR).	1	1
266	U27269	Mus	sodium glucose cotransporter	204	56
	1	musculus		1	
267	AF124491	Homo sapiens	ARF GTPase-activating protein GIT2	159	75
268	AF127389	Rattus	putative taste receptor TR1	209	39
		norvegicus	,		
269	X98296	Homo sapiens	ubiquitin hydrolase	215	95
270	X78482	Streptococcus	Fo-gamma receptor	129	26
	1	pyogenes	T T Barrella T T T T T T T T T T T T T T T T T T	1.2	1 -0
271	AB009883	Nicotiana	KED	109	26
		tabacum			
272	AF137367	Mus	VPS10 domain receptor protein SORCS	899	97
	1	musculus	The desired process process of the p	""	1
273	L34938	Rattus	ionotropic glutamate receptor	460	86
	254750	norvegicus	restoropie glacamate receptor	1 700	60
274	AL022724	Homo sapiens	dJ413H6.1.1 (hamster Androgen-dependent	188	74
		Tromo suprems	Expressed Protein LIKE PUTATIVE protein)	1 .00	'
	ı	İ	(isoform 1)	1	1
275	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme	173	94
213	Al 203333	Homo Sapiens	APOLLON	1/3	34
276	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	148	56
277					
	L40380	Homo sapiens	thyroid receptor interactor	430	61
278	AB046851	Homo sapiens	KIAA1631 protein	283	96
279	AC008075	Arabidopsis	Contains PF 00069 Eukaryotic protein kinase	157	43
202	1 (02722	thaliana	domain.	 	
280	M83738	Homo sapiens	protein-tyrosine phosphatase	181	73
281	AK024397	Homo sapiens	unnamed protein product	439	91
282	AF141326	Homo sapiens	RNA helicase HDB/DICE1	197	84
283	AF156530	Mus	ETS-domain transcriptional repressor PE1	605	76
	<u> </u>	musculus		<u> </u>	
284	Y29336	Homo sapiens	Human secreted protein clone cs756_2 alternate	647	100
			reading frame protein.	1	
285	Y73402	Homo sapiens	Human secreted protein clone yc25_1 protein	300	90
		1	sequence SEQ ID NO:26.	1	1
	AF016411	Homo sapiens	KCNA3.1B	137	100
286	120.0111				
286 287	W89253	Homo sapiens	Human ALP.	688	97
		Homo sapiens Bos taurus		750	
287	W89253 AF112886	Bos taurus	differentiation enhancing factor 1	750	96
287 288	W89253				96

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:				Score	Identity
	17105054	norvegicus			
292	AF102854	Rattus norvegicus	membrane-associated guanylate kinase- interacting protein 2 Maguin-2	124	53
293	X99211	Drosophila melanogaster	ubiquitin-specific protease	143	38
294	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	185	94
295	Y94890	Homo sapiens	Human protein clone HP02798.	108	59
296	AF019767	Homo sapiens	zinc finger protein	154	96
297	Y28568	Homo sapiens	Secreted peptide clone bd577 1.	568	84
298	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	182	97
299	B08906	Homo sapiens	Human secreted protein sequence encoded by gene 16 SEQ ID NO:63.	605	69
300	R58890	Homo sapiens	Human-32 cadherin-related molecule.	212	97
301	AF022859	Homo sapiens	neuropilin-2(a0)	277	100
302	Y71124	Homo sapiens	Human mitogenic regulator duox2.	716	97
303	Y44297	Homo sapiens	Human receptor tyrosine kinase.	228	97
304	D32050	Homo sapiens	alanyi-tRNA synthetase	192	80
305	U43586	Homo sapiens	protein kinase related to Raf protein kinases; Method: conceptual translation supplied by author	428	72
306	R54872	Homo sapiens	Human H13 viral receptor mutant 4.	280	95
307	D78572	Mus musculus	membrane glycoprotein	199	41
308	AF255614	Rattus norvegicus	scaffolding protein SLIPR	639	88
309	S79463	Mus sp.	semaphorin homolog=M-Sema F	162	89
310	AF178941	Homo sapiens	ATP-binding cassette sub-family A member 2	736	100
311	U03413	Dictyostelium discoideum	calcium binding protein	151	36
312	Y87347	Homo sapiens	Human signal peptide containing protein HSPP- 124 SEQ ID NO:124.	744	100
313	Z97055	Homo sapiens	dJ388M5.4 (putative GS2 like protein)	789 .	99
314	AC004010	Homo sapiens	similar to Leucine-rich transmembrane proteins; 44% similarity to U42767 (PID:g1736918)	197	38
315	AL021392	Homo sapiens	dJ439F8.2 (supported by GENSCAN and GENEWISE)	278	38
316	U70209	Mus musculus	polycystic kidney disease I protein	165	38
317	AF109643	Rattus norvegicus	coxsackie-adenovirus-receptor homolog	223	38
318	AF104923	Homo sapiens	putative transcription factor	138	84
319	AF100287	Trypanosoma vivax	activated protein kinase C receptor homolog	141	38
320	G00588	Homo sapiens	Human secreted protein, SEQ ID NO: 4669.	125	51
321	Y21591	Homo sapiens	Human secreted protein (clone CC332-33).	459	97
322	D26070	Homo sapiens	human type 1 inositol 1,4,5-trisphosphate receptor	232	97
323	Y27918	Homo sapiens	Human secreted protein encoded by gene No. 123.	306	88
324	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	209	70
325	M19650	Homo sapiens	2',3'-cyclic-nucleotide 3'-phosphodiesterase (EC 3.1.4.37)	214	97
326	W80396	Homo sapiens	A secreted protein encoded by clone bp646_10.	140	70
327	X75756	Homo sapiens	protein kinase C mu	540	78
328	G02292	Homo sapiens	Human secreted protein, SEQ ID NO: 6373.	721	99
329	AF168990	Homo sapiens	putative GTP-binding protein	877	99
330	S67984	Homo sapiens	anti-HIV gp120 antibody heavy chain variable region	581	80
331	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	2823	98
332	Y87330	Homo sapiens	Human signal peptide containing protein HSPP- 107 SEQ ID NO:107.	1127	100
333	Y28503	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	320	98
334	AC002563	Homo sapiens	putative RHO/RAC effector protein; 95%	327	93

SEQ	Accession	Species	Description	Smith-	%
ID `	No.	1	·	Waterman	Identity
NO:	1		_	Score	
			similarity to P49205 (PID:g1345860)		
335	Y87347	Homo sapiens	Human signal peptide containing protein HSPP-	1111	67
		.L	124 SEQ ID NO:124.		<u> </u>
336	AF006466	Mus musculus	lymphocyte specific formin related protein	193	75
337	AF265555	Homo sapiens	ubiquitin-conjugating BIR-domain enzyme APOLLON	632	97
338	Y13443	Homo sapiens	Amino acid sequence of hSlo3-2.	516	100
339	Y07637	Homo sapiens	putative GABA-gated chloride channel	189	100
340	Y05734	Homo sapiens	Human Grb7 effector 2.2412 protein.	2156	99
341	AE000497	Escherichia coli	L-idonate transcriptional regulator	928	98
342	D90855	Escherichia coli	glycerol-3-phosphate dehydrogenase (EC 1.1.99.5) chain A, anaerobic	769	99
343	D85613	Escherichia coli	membrane component	399	100
344	M93239	Escherichia	transmembrane protein	232	100
345	M60177	coli Escherichia	enterobactin	759	99
		coli			
346	D90699	Escherichia coli	Sensor protein copS (EC 2.7.3).	638	97
347	D90843	Escherichia coli	CapB protein.	552	100
348	M13422	Escherichia coli	49 kd protein	1193	96
349	L10328	Escherichia coli	similar to drug resistance translocases	340	90
350	X69942	Mus musculus	enhancer-trap-locus-1	560	82
351	AF239613	Homo sapiens	apamin-sensitive small-conductance Ca2+-	463	80
250	2000000	<u> </u>	activated potassium channel	577	100
352	D90777	Escherichia coli	3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157) (b- hydroxybutyryl-CoA dehydrogenase) (BhbD).	377	100
353	D90863	Escherichia coli	similar to	311	98
354	Y52386	Homo sapiens	Human transmembrane protein HP02000.	133	58
355	Y31645	Homo sapiens	Human transport-associated protein-7 (TRANP-7).	482	55
356	Y58637	Homo sapiens	Protein regulating gene expression PRGE-30.	119	51
357	AF119226	Homo sapiens	dual-specificity tyrosine phosphatase YVHI	1788	100
358	Y87219	Homo sapiens	Human secreted protein sequence SEQ ID NO:258.	165	100
359	J00132	Homo sapiens	beta-fibrinogen	233	93
360	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	128	70
361	R28916	Homo sapiens	Type III procollagen (prior art).	108	40
362	U16655	Rattus norvegicus	phospholipase C delta-4	649	65
363	G03119	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.	95	42
364 -	U47276	Gallus gallus	chicken brain factor-2	104	34
365	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	183	65
366	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	118	46
367	X98258	Homo sapiens	M-phase phosphoprotein 9	564	75
368	AL021366	Homo sapiens	clCK0721Q.3 (Kinesin related protein)	3387	99
369	U70932	Peromyscus leucopus	reverse transcriptase	92	59
370	X86400	Homo sapiens	gamma subunit of sodium potassium ATPase like	242	73
371	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	165	56
372	U49974	Homo sapiens	mariner transposase	257	55
373	X13916	Homo sapiens	LDL-receptor related precursor (AA -19 to 4525)	21 193	99
374	AF234765	Rattus norvegicus	serine-arginine-rich splicing regulatory protein SRRP86	1182	78
375	U49974	Homo sapiens	mariner transposase	172 .	55

SEQ	Accession	Species	Description	Smith-	1%
ID ID	No.	эрсско	Description	Waterman	Identity
NO:	1	'		Score	Ideality
376	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	221	67
377	G00669	Homo sapiens	Human secreted protein, SEQ ID NO: 4750.	600	100
378	X52574	Mus	GTP binding protein	1456	91
		musculus		1	1
379	R69095	Homo sapiens	Anti-HIV Fab tat31 light chain.	68	37
380	J04974	Homo sapiens	alpha-2 type XI collagen	125	37
381	AB002405	Homo sapiens	LAK-4p	530	43
382	U64830	Dictyostelium	protein tyrosine kinase	115	44
	}	discoideum		1	
383	G02916	Homo sapiens	Human secreted protein, SEQ ID NO: 6997.	618	98
384	G01194	Homo sapiens	Human secreted protein, SEQ ID NO: 5275.	617	93
385	AJ245822	Homo sapiens	type I transmembrane receptor	4560	100
386	D86974	Homo sapiens	KIAA0220	2148	98
387	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	142	50
388	G04072	Homo sapiens	Human secreted protein, SEQ ID NO: 8153.	99	59
389	M12140	Homo sapiens	envelope protein	197	51
390	AJ293309	Homo sapiens	NHP2 protein	461	77
391	Y42751	Homo sapiens	Human calcium binding protein 2 (CaBP-2).	181	94
392	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	241	66
393	Y14442	Homo sapiens	olfactory receptor protein	339	54
394	W85607	Homo sapiens	Secreted protein clone da228_6.	957	100
395	Y76332	Homo sapiens	Fragment of human secreted protein encoded by gene 38.	171	34
396	G03930	Homo sapiens	Human secreted protein, SEQ ID NO: 8011.	250	100
397	AB032904	Hylobates syndactylus	dopamine receptor D4	105	35
398	AJ007798	Homo sapiens	stromal antigen 3, (STAG3)	861	.85
399	Y91405	Homo sapiens	Human secreted protein sequence encoded by gene 2 SEQ ID NO:126.	1047	92
400	Y29861	Homo sapiens	Human secreted protein clone cb98 4.	162	37
401	D87002	Homo sapiens	similar to rat integral membrane glycoprotein; accession number Z21513.	527	78
402	AF100754	Homo sapiens	ancient ubiquitous protein AUP1 isoform	853	95
403	X74904	Gallus gallus	alpha-2-macroglobulin receptor	258	60
404	AF075462	Mus musculus	ADP-ribosylation factor-directed GTPase activating protein isoform b	545	89
405	X92887	Human endogenous retrovirus K	pol/env	162	30
406	Y30162	Homo sapiens	Human dorsal root receptor 4 hDRR4.	325	72
407	AK022626	Homo sapiens	unnamed protein product	2833	99
408	L13802	Homo sapiens	ribosmal protein small subunit	264	92
409	Y91600	Homo sapiens	Human secreted protein sequence encoded by gene 9 SEQ ID NO:273.	1788	89
410	W88745	Homo sapiens	Secreted protein encoded by gene 30 clone HTSEV09.	2004	99
411	AB043953	Mus musculus	Chat-H	2628	82
412	Y86233	Homo sapiens	Human secreted protein HNTMX29, SEQ ID NO:148.	1014	92
413	U10542	Pan troglodytes	MHC class I A	265	71
414	AF155097	Homo sapiens	NY-REN-7 antigen	850	95
415	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	88	48
416	Y57911	Homo sapiens	Human transmembrane protein HTMPN-35.	266	89
417	W27651	Homo sapiens	Secreted protein AT205.	481	60
418	Y76884	Homo sapiens	Retinoblastoma binding protein-7sequence.	3077	87
419	AF255559	Notothenia coriiceps	alpha tubulin	289	68
420	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	209	74
421	AL109827	Homo sapiens	dJ309K20.2 (acrosomal protein ACR55 (similar to rat sperm antigen 4 (SPAG4)))	1446	96
422	AC008075	Arabidopsis	F24J5.4	112	35

SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:		<u> </u>		Score	100
423	AF231705	Homo sapiens	Alu co-repressor 1	1090	100
424	AF234887	Homo sapiens	FLAMINGO I	6268 1961	97
425	Y35942	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 191.	1901	99
426	AB009288	Homo sapiens	N-copine	635	98
420	L12392	Homo sapiens	Huntington's Disease protein	16080	99
428	Y94990	Homo sapiens	Human secreted protein vb21_1, SEQ ID NO:20.	768	98
429	AJ293573	Homo sapiens	zinc finger protein Cezanne	542	87
430	Y84441	Homo sapiens	Amino acid sequence of a human RNA-	2074	100
			associated protein.	ł	1
431	G02850	Homo sapiens	Human secreted protein, SEQ ID NO: 6931.	723	95
432	G04067	Homo sapiens	Human secreted protein, SEQ ID NO: 8148.	73	42
433	AF159296	Lycopersicon	extensin-like protein	613	48
		esculentum			
434	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	135	44
435	X73874	Homo sapiens	phosphorylase kinase	3442	97
436	AF161426	Homo sapiens	HSPC308	268	74
437	Y30812	Homo sapiens	Human secreted protein encoded from gene 2.	1055	52 i 56
438	G03798 X14766	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7879. GABA-A receptor alpha 1 subunit	2294	96
440	X02344	Homo sapiens	beta-tubulin	311	95
441	AF168418	Homo sapiens	activating signal cointegrator 1	1882	100
442	L11672	Homo sapiens	zinc finger protein	795	1 54
443	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	93	26
444	A52140	unidentified	HUMAN NDR	2451	100
445	X98330	Homo sapiens	ryanodine receptor 2	9356	99
446	AF116712	Homo sapiens	PRO2738	227	49
447	AF245447	Homo sapiens	sphingosine kinase type 2 isoform	576	99
448	AF133086	Homo sapiens	membrane-type serine protease 1	2630	94
449	U87305	Rattus	transmembrane receptor UNC5H1	817	93
		norvegicus			
450	AF081249	Homo sapiens	JAW1-related protein MRVIIA long isoform	4568	99
451	AC005498	Homo sapiens	R31665_1	316	62
452	M60235	Homo sapions	granule membrane protein-140	730	73 88
453 454	AB036706 G00918	Homo sapiens Homo sapiens	intelectin Human secreted protein, SEQ ID NO: 4999.	263	81
455	Y22634	Homo sapiens	Human cytokine inducible regulatory protein-1	192	67
422	122034	riomo sapiens	(CIRP-1).	172	"
456	Y36705	Homo sapiens	Fragment of human secreted protein encoded by	106	40
	1		gene 62.		[
457	N91325	Homo sapiens	DNA encoding human growth hormone receptor.	3282	96
458	M19155	Plasmodium	S-antigen precursor	110	36
		falciparum			<u></u>
459	Y13377	Homo sapiens	Amino acid sequence of protein PRO257.	509	98
460	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	149	43
	ļ		clone HTDAD22.		
461	Y14482	Homo sapiens	Fragment of human secreted protein encoded by	184	54
463	V62006	Homo sapiens	gene 17.	135	47
462	Y53005	Homo sapiens	Human secreted protein clone pm749_8 protein sequence SEO ID NO:16.	133	4'
463	X84960	Triticum	low molecular weight glutenin	109	33
403	1.01500	aestivum	i iow molecular weight gratemin	'*'	""
464	W19919	Homo sapiens	Human Ksr-1 (kinase suppressor of Ras).	1781	85
465	AF189764	Mus	alpha/beta hydrolase-i	502	59
	1	musculus	1	1	
466	U93569	Homo sapiens	p40	101	30
467	Y41528	Homo sapiens	Fragment of human secreted protein encoded by	1172	99
			gene 77.		İ
468	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	149	52
469	AJ000008	Homo sapiens	PI3-kinase	5832	97
470	X70922	Mus	neurotoxin homologue	118	47
470			1	1	1
471	G03797	musculus Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	198	75

SEQ ID	Accession No.	Species	Description	Smith- Waterman Score	% Identity
10:				Score	
			gene 62.	328	100
173	G02313	Homo sapiens	Human secreted protein, SEQ ID NO: 6394.	1013	97
174	¥07007	Homo sapiens	Breast cancer associated antigen precursor sequence.		1
475	W93254	Homo sapiens	Human ESRP1 protein.	943	80
476	W48351	Homo sapiens	Human breast cancer related protein BCRB2.	236	65
477	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	202	60
478	G01870	Homo sapiens	Human secreted protein, SEQ ID NO: 5951.	267	100
479	AF102777	Mus musculus	FYVE finger-containing phosphoinositide kinase	3427	92
400	C02052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	123	53
480	G03052		A human membrane fusion protein designated	221	77
481	W87701	Homo sapiens	SYTAX1.	131	39
482	G03119	Homo sapiens	Human secreted protein, SEQ ID NO: 7200.		59
483	AF210651	Homo sapiens	NAG18	124	50
484	AF010144	Homo sapiens	neuronal thread protein AD7c-NTP	343	70
485	G00637	Homo sapiens	Human secreted protein, SEQ ID NO: 4718.	129	
486	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein 3	149	73
487	Y76167	Homo sapiens	Human secreted protein encoded by gene 44.	627	100
488	AJ275213	Homo sapiens	stabilin-1	1244	91
489	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	313	65
490	L12392	Homo sapiens	Huntington's Disease protein	16081	100
491	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	197	66
492	J03799	Homo sapiens	laminin-binding protein	228	70
493	U15174	Homo sapiens	BCL2/adenovirus E1B 19kD-interacting protein	128	41
494	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	197	67
405	AC005175	Homo sapiens	R31449 3	889	94
495		Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	229	61
496	G03786	Canis	D4 dopamine receptor	90	48
497	AB030237	familiaris			65
498	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	228	52
499	U70935	Peromyscus maniculatus	reverse transcriptase	213	
500	U48508	Homo sapiens	skeletal muscle ryanodine receptor	26406	99
501	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	105	58
502	AF119851	Homo sapiens	PRO1722	156	62
503	AF113685	Homo sapiens	PRO0974	116	50
504	U79458	Homo sapiens	WW domain binding protein-2	322	59
505	W29651	Homo sapiens	Human secreted protein CD124 3.	608	55
506	W85459	Homo sapiens	Secreted protein encoded by clone dhi 135 9.	986	70
507	Y86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID NO:180.	115	33
508	AL160175	Homo sapiens	bA243J16.3 (similar to MYLK (myosin, light polypeptide kinase))	184	92
509	U43360	Peromyscus maniculatus	reverse transcriptase	97	62
610	C02700	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	1,17	63
510	G03789	Homo sapiens		1058	100
511	W79092	Homo sapiens		205	64
512	AF010144	Homo sapiens		2151	100
513	AJ133439		CG6393 gene product		42
514	AE003456	Drosophila melanogaster		237	
515	Z17206	Xenopus laevis	p46XlEg22	128	40
516	AF104413	Homo sapiens	large tumor suppressor 1	1766	94
517	G03797	Homo sapiens		92	40
518		Homo sapiens	HSPC249	444	98
519	S80864	Homo sapiens		318	50
520	X92485	Plasmodium vivax	pval	170	61

SEQ	Accession	Species	Description	Smith-	%
ID `	No.	'	•	Waterman	Identity
NO:	i	i		Score	1
569	AF097518	Homo sapiens	liver-specific transporter	231	100
570	AB035698	Homo sapiens	Misshapen/NIK-related kinase MINK-1	1532	100
571	Y07096	Homo sapiens	Colon cancer associated antigen precursor	1064	100 .
	1		sequence.		
572	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	735	55
573	Y66639	Homo sapiens	Membrane-bound protein PRO290.	254	45
574	AB037108	Homo sapiens	seven transmembrane domain orphan receptor	1883	99
575	D43949	Homo sapiens	This gene is novel.	836	100
576	Y48596	Homo sapiens	Human breast tumour-associated protein 57.	108	50
577	G00352	Homo sapiens	Human secreted protein, SEQ ID NO: 4433.	141	75
578	R95913	Homo sapiens	Neural thread protein.	140	65
			unnamed protein product	201	70
579	AK025116	Homo sapiens			
580	Y86473	Homo sapiens	Human gene 52-encoded protein fragment, SEQ ID NO:388.	77	70
581	AF:196779	Homo sapiens	JM10 protein	450	100
582	AF188706	Homo sapiens	g20 protein	330	98
583	AB030234	Canis	D4 dopamine receptor	64	56
-601	00000	familiaris	V	1345	000
584	G02621	Homo sapiens	Human secreted protein, SEQ ID NO: 6702.	345	90
585	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2 Antigen)	268	85
586	Y30819	Homo sapiens	Human secreted protein encoded from gene 9.	235	35
587	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	132	56
588	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	182	79
589	AF235017	Mus	2P1 protein	764	80
590	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	329	81
591	Y30709	Homo sapiens	Amino acid sequence of a human secreted protein.	110	43
592	Y53875	Homo sapiens	A human seven transmembrane signal transducer	1369	92
593	Y53051	Homo sapiens	polypeptide. Human secreted protein clone dd119_4 protein	1112	97
		•	sequence SEQ ID NO:108.		
594	Y27658	Homo sapiens	Human secreted protein encoded by gene No. 92.	763	79
595	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 7879.	156	58
596	AF151110	Mus musculus	COPI protein	2215	95
597	G03786	Homo sapiens	Human secreted protein, SEQ ID NO: 7867.	157	65
598	AF192499	Mus	putative secreted protein ZSIG37	143	40
	<u> </u>	musculus		L	<u> </u>
599	AF119855	Homo sapiens	PRO1847	236	76
600·	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	212	73
601	Y00295	Homo sapiens	Human secreted protein encoded by gene 38.	567	88
602	AF184971	Homo sapiens	class II cytokine receptor ZCYTOR7	2015	74
603	AF061936	Homo sapiens	diacylglycerol kinase iota	773	96
604	AL096828	Homo sapiens	dJ963E22.1 (Novel protein similar to NY-REN-2 Antigen)	1333	93
605	AB033106	Homo sapiens	KIAA1280 protein	3915	100
606	X75756	Homo sapiens	protein kinase C mu	3916	99
607	D86983	Homo sapiens	similar to D.melanogaster peroxidasin(U11052)	5758	99
608	W69341	Homo sapiens		1377	99
609			Secreted protein of clone CG279_1.		
	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	339	82
610	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	116	62
611	AF202636	Homo sapiens	angiopoietin-like protein PP1158	2164	100
612	AF090944	Homo sapiens	PRO0663	218	82
613	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	195	59
614	M87053	Rattus	lens membrane protein	450	84
	1.0001000	norvegicus		L	
615	AC004232	Homo sapiens	FPM315	163	37
616	G01984	Homo sapiens	Human secreted protein, SEQ ID NO: 6065.	205	79

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
617	Y91524	Homo sapiens	Human secreted protein sequence encoded by gene 74 SEQ ID NO:197.	821	99
618	AJ245621	Homo sapiens	CTL2 protein	2258	99
619	Y76198	Homo sapiens	Human secreted protein encoded by gene 75.	108	64
620	AF067864	Homo sapiens	transferrin receptor 2 alpha	3922	94
621	D90721	Escherichia coli	Transmembrane protein dppC	573	90
622	W75858	Homo sapiens	Human secretory protein of clone CS752-3.	730	100
623	Y94982	Homo sapiens	Human secreted protein vb12_1, SEQ ID NO:4.	733	100
624	AF034745	Mus musculus	LNXp80	637	83
625	U42580	Paramecium bursaria Chlorella virus 1	Pro-rich, IPPPNMSLPLS (3x)	94	46
626	U79260	Homo sapiens	unknown	194	70
627	R95913	Homo sapiens	Neural thread protein.	99	50
628	G03450	Homo sapiens	Human secreted protein, SEQ ID NO: 7531.	427	100
629	Y36281	Homo sapiens	Human secreted protein encoded by gene 58.	590	100
630	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	165	76
631	G02139	Homo sapiens	Human secreted protein, SEQ ID NO: 6220.	268	96
632	U16996	Homo sapiens	protein tyrosine posphatase	351	80
633	AF121857	Homo sapiens	sorting nexin 7	2019	100
634	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10 encoded by GenBank Accession Number L25899	340	77
635	Y07090	Homo sapiens	Renal cancer associated antigen precursor sequence.	277	64
636	AB013382	Homo sapiens	DUSP6	414	76
637	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	315	71
638	M95762	Rattus norvegicus	GABA transporter	924	89
639	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	219	60
640	Ý01400	Homo sapiens	Secreted protein encoded by gene 18 clone HNHFO29.	137	79
641	AC008075	Arabidopsis thaliana	F24J5.4	121	33
642	W74824	Homo sapiens	Human secreted protein encoded by gene 96 clone HAQBK61.	615	62
643	AB015982	Homo sapiens	serine/threonine kinase	485	98
644	Y25806	Homo sapiens	Human secreted protein fragment encoded from gene 23.	162	46
645	AF122904	Homo sapiens	membrane protein DAP10	474	100
646	AF233323	Homo sapiens	Fas-associated phosphatase-1	200	38
647	W48804	Homo sapiens	Homo sapiens clone BK158_1 protein.	1203	99
648	AF257330	Homo sapiens	COBW-like protein	1440	98
649	Y36203	Homo sapiens	Human secreted protein #75.	233	73
650 651	G02872 Y32199	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 6953. Human receptor molecule (REC) encoded by	173	78 100
652	AB032909	Hylobates	Incyte clone 2022379. dopamine receptor D4	122	32
653	AK021848	agilis	Name of the last o		-
654	W73411	Homo sapiens Homo sapiens	unnamed protein product	186	69
655	L22455	Rattus	Human secreted protein encoded by Gene No. 15.	57	37
	· · ·	norvegicus	mu opioid receptor	116	34
656	G03112	Homo sapiens	Human secreted protein, SEQ ID NO: 7193.	110	45
657 658	G02345	Homo sapiens	Human secreted protein, SEQ ID NO: 6426.	459	97
038	W88627	Homo sapiens	Secreted protein encoded by gene 94 clone HPMBQ32.	291	75
	G00000	+++			
659 660	G02832 Y91423	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 6913. Human secreted protein sequence encoded by	134 333	65 96

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
661	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	168	68
662	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	375	43
663	W75771	Homo sapiens	Human GTP binding protein APD08.	629	100
664	AL096770	Homo sapiens	bA150A6.2 (novel 7 transmembrane receptor	480	55
			(rhodopsin family) (olfactory receptor like) protein (hs6M1-21))		
665	AB037734	Homo sapiens	KIAA1313 protein	978	96
666	W82841	Homo sapiens	Human cerebral protein-1.	192	84
667	W82841	Homo sapiens	Human cerebral protein-1.	182	87
668	AB030184	Mus musculus	contains transmembrane (TM) region and ATP binding region	757	68
669	AB032919	Hylobates muelleri	dopamine receptor D4	85	37
670	AF107295	Rattus norvegicus	outer membrane protein	746	81
671	Z33642	Homo sapiens	leukocyte surface protein	394	93
672	W85608	Homo sapiens	Secreted protein clone du410 5.	261	91
673	G03203	Homo sapiens	Human secreted protein, SEQ ID NO: 7284.	106	48
674	AL035587	Homo sapiens	dJ475N16.4 (KIAA0240)	2388	99
675	Y59668	Homo sapiens	Secreted protein 108-005-5-0-C1-FL.	1134	53
676	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	174	74
677	AF026954	Bos taurus	pyruvate dehydrogenase phosphatase regulatory subunit precursor; PDPr	1013	95
678	L11625	Mus musculus	receptor protein-tyrosine kinase	545	96
679	AL031427	Homo sapiens	dJ167A19.3 (novel protein)	745	100
680	AJ133430	Mus musculus	olfactory receptor	528	77
681	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	179	70
682	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	336	76
683	Y94943	Homo sapiens	Human secreted protein clone yt14_1 protein sequence SEQ ID NO:92.	118	100
684	U43360	Peromyscus maniculatus	reverse transcriptase	100	37
685	G00885	Homo sapiens	Human secreted protein, SEQ ID NO: 4966.	162	60
686	AK001518	Homo sapiens	unnamed protein product	590	100
687	G01982	Homo sapiens	Human secreted protein, SEQ ID NO: 6063.	718	100
688	Y92241	Homo sapiens	Human cancer associated antigen precursor (MO-REN-46).	2405	99
689	AC024792	Caenorhabditi s elegans	contains similarity to TR:P78316	423	36
690	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	183	81
691	Y56514	Homo sapiens	Human Jurkat cell clone P2-15 AIM10 longest ORF protein sequence.	180	88 .
692	Y27795	Homo sapiens	Human secreted protein encoded by gene No. 79.	1539	99
693	Y36268	Homo sapiens	Human secreted protein encoded by gene 45.	428	98
694	U12465	Homo sapiens	ribosomal protein L35	308	89
695	Y45272	Homo sapiens	Human secreted protein encoded from gene 16.	1517	99
696	AF191838	Homo sapiens	TANK binding kinase TBK1	1242	98
697	Y02693	Homo sapiens	Human secreted protein encoded by gene 44 clone HTDAD22.	275	75
698	Y87280	Homo sapiens	Human signal peptide containing protein HSPP- .57 SEQ ID NO:57.	576	90
699	Y97999	Homo sapiens	Human SCAD family molecule HSFM-1, SEQ ID NO:1.	729	99
700	AJ006701	Homo sapiens	putative serine/threonine protein kinase	610	79
701	AF209198	Homo sapiens	zinc finger protein 277	2357	100
702	AJ298841	Mus musculus	torsinA protein	709	45
703	AK021729	Homo sapiens	unnamed protein product	622	98
704	Z46787	Caenorhabditi s elegans	similar to Glutaredoxin, Zinc finger, C3HC4 type (RING finger)	920	51
705	G02882	Homo sapiens	Human secreted protein, SEQ ID NO: 6963.	589	98

050				15 10	1
SEQ	Accession	Species	Description	Smith-	%
ID	No.			Waterman	Identity
NO:	-	 		Score	<u> </u>
706	G02501	Homo sapiens	Human secreted protein, SEQ ID NO: 6582.	125	58
707	R95326	Homo sapiens	Tumor necrosis factor receptor 1 death domain	121	95
	<u> </u>	<u> </u>	ligand (clone 2DD).		
708	G03002	Homo sapiens	Human secreted protein, SEQ ID NO: 7083.	125	39
709	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	516	98
710	M63577	Saccharomyc	SFP1	131	59
		es cerevisiae		ļ	
711	AB026291	Rattus	acetoacetyl-CoA synthetase	467	85
	i	norvegicus			1
712	D21211	Homo sapiens	protein tyrosine phosphatase (PTP-BAS, type 3)	368	44
713	AF044033	Marmota	olfactory receptor	615	83
	1	marmota		1	-
714	G03561	Homo sapiens	Human secreted protein, SEQ ID NO: 7642.	251	100
715	AB033062	Homo sapiens	KIAA1236 protein	1380	100
716	G00577	Homo sapiens	Human secreted protein, SEQ ID NO: 4658.	80	73
717	Y96864	Homo sapiens	SEQ. ID. 37 from WO0034474.	835	99
718	AJ243396	Homo sapiens	voltage-gated sodium channel beta-3 subunit	234	100
719	U47334	Homo sapiens	similar to chicken gamma aminobutyric acid	578	99
-			receptor beta4 subunit	1	
720	AB020598	Homo sapiens	peptide transporter 3	1096	100
721	Y53886	Homo sapiens	A suppressor of cytokine signalling protein	570	74
		. Supreits	designated HSCOP-6.	1 3/2	'-
722	J05046	Homo sapiens	insulin receptor-related receptor	6787	100
723	AF001958	Ambystoma	electrogenic Na+ bicarbonate cotransporter;	111	41
, 20	111 001550	tigrinum	NBC	1 ***	1 * '
724	AF127084	Mus	semaphorin cytoplasmic domain-associated	5253	94
	1.2.2700	musculus	protein 3A	3233) 74
725	X54673	Homo sapiens	GABA transporter	3114	99
726	AF016191	Rattus	potassium channel	370	100
720	74 010171	norvegicus	potassium chainer	1 370	100
727	AB029559	Rattus	BATI	139	35
121	ABOZE	norvegicus	DA14	139	33
728	Y28503	Homo sapiens	HGFH3 Human Growth Factor Homologue 3.	2186	97
729	AJ011415	Homo sapiens	plexin-B1/SEP receptor	729	56
730	Z93096	Homo sapiens	bK390B3.1 (manic fringe (Drosophila)	142	68
750	2,530,90	nomo sapiens	homolog)	142	08
731	Z10062	Homo sapiens	cDNA encoding a human vanilloid receptor	675	00
/31	210002	nomo sapiens	homologue Vanilrep1.	0/3	99
732	AF161382	Homo sapiens	HSPC264	492	94
733	AB029033	Homo sapiens	KIAA1110 protein	3826	1
734	AE000493	Escherichia	KIAAT I to protein		99
/34	AE000493	coli	putative transport protein	592	97
735	AL033379	Homo sapiens	47417022.2 () 7	2173	99
133	AL033379	riomo sapiens	dJ417O22.2 (novel 7 transmembrane receptor	21/3	99
	ł	1	(rhodopsin family) protein similar to high-	ł	ì
736	AF132599	Homo sapiens	affinity lysophosphatidic acid receptor homolog)	245	-
130	אניייייייייייי	Tromo sabiens	RANTES factor of late activated T lymphocytes-	245	56
737	X55019	Homo sapiens	nastriebalina maantan dalta	002	100
738	X91906	Homo sapiens	acetylcholine receptor delta subunit	883	99
739			voltage-gated chloride ion channel	1978	100
740	AB026116	Homo sapiens	organic anion transporter 4	1444	98
740	D00570	Mus	open reading frame (196 AA)	83	24
741	11/02/26	musculus	TY	<u> </u>	ļ
	W03626	Homo sapiens	Human thyrotropin GPR N-terminal sequence.	118	40
742	U66059	Homo sapiens	V_segment translation product	614	100
743	AF119815	Homo sapiens	G-protein-coupled receptor	2751	99
744	X16663	Homo sapiens	haematopoietic lineage cell protein (AA 1-486)	148	93
745	W67838	Homo sapiens	Human secreted protein encoded by gene 32	448	95
	L		clone HLTCJ63.		
746	W57260	Homo sapiens	Human semaphorin Y.	2414	100
747	W21578	Homo sapiens	Alzheimer's disease protein encoded by DNA	968	65
/4/	11.21370	1 .			
	L		from plasmid pGCS2232.		L
748	Y94935	Homo sapiens		622	100
748	Y94935	1	from plasmid pGCS2232. Human secreted protein clone yd218_1 protein sequence SEQ ID NO:76.	622	100
	L	Homo sapiens Homo sapiens Homo sapiens	from plasmid pGCS2232. Human secreted protein clone yd218_1 protein	622	100

SEQ ID	Accession No.	Species	Description	Smith- Waterman Score	% Identity
NO: 751	AB025258	Mus	granuphilin-a	773	41
		musculus			
752	Y52386	Homo sapiens	Human transmembrane protein HP02000.	900	99
753	Y48586	Homo sapiens	Human breast tumour-associated protein 47.	2527	99
754	AJ272207	Homo sapiens	putative G protein-coupled receptor 92	694	100
755	M85183	Rattus	vasopressin receptor	979	68
756	AF190501	norvegicus Homo sapiens	leucine-rich repeat-containing G protein-coupled	388	71
757	Y02692	Homo sapiens	Human secreted protein encoded by gene 43	461	87
750	700626		clone HTADX17.	120	00
758 759	Z22535 R04932	Homo sapiens Homo sapiens	ALK-3 Interferon-gamma receptor segment from clone 39 responsible for binding the target.	439 564	98 97
760	W74902	Homo sapiens	Human secreted protein encoded by gene 175 clone HE8BI92.	1217	99
761	G03706	Homo sapiens	Human secreted protein, SEQ ID NO: 7787.	223	88
762	AB020676	Homo sapiens	KIAA0869 protein	4433	99
763	AK026992	Homo sapiens	unnamed protein product	2285	99
764	AF173358	Homo sapiens	glucocorticoid receptor AF-1 coactivator-1	573	100
765	AF268066	Mus	netrin 4	2019	89
766	Y48585	Homo sapiens	Human breast tumour-associated protein 46.	1169	89
767	AF230378	Mus	interleukin-l delta	309	45
768	AF121975	musculus Mus	odorant receptor S18	268	62
=	150000	musculus			<u> </u>
769	AB008515	Homo sapiens	RanBPM	611	57
770	Y09945	Rattus norvegicus	putative integral membrane transport protein	458	50
771	AF226731	Homo sapiens	AD026	688	99
772	Y27132	Homo sapiens	Human glioblastoma-derived polypeptide (clone OA004FG).	1384	100
773	X87832	Homo sapiens	NOV/plexin-A1 protein	1821	98
774	AB025258	Mus musculus	granuphilin-a	500	41
775	AF125101	Homo sapiens	HSPC040 protein	232	93
776	G02815	Homo sapiens	Human secreted protein, SEQ ID NO: 6896.	314	95
777	G02493	Homo sapiens	Human secreted protein, SEQ ID NO: 6574.	191	68
778	R03301	Homo sapiens	Sequence of pre-human atrial natriuretic peptide.	213	45
779	AL357374	Homo sapiens	bA353C18.2 (novel protein)	232	100
780	AF100346	Homo sapiens	neuronal voltage gated calcium channel gamma- 3 subunit	1434	89
781	Y19566	Homo sapiens	Amino acid sequence of a human secreted protein.	103	52
782	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	1098	93
783	AF084464	Rattus	GTP-binding protein REM2	141	30
784	W49042	norvegicus Homo sapiens	Human low density lipoprotein binding protein LBP-3.	2693	99
785	AF238381	Homo sapiens	PTOVI	1904	91
786	Y91870	Homo sapiens	Human apoptosis related protein.	547	100
787	Y71062	Homo sapiens	Human membrane transport protein, MTRP-7.	1062	94
788	AF117754	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP240	8684	98
789	AL049569	Homo sapiens	dJ37C10.3 (novel ATPase)	2848	96
790	AF151848	Homo sapiens	CGI-90 protein	745	96
791	Y08639	Homo sapiens	nuclear orphan receptor ROR-beta	1421	95
792	Y41706	Homo sapiens	Human PRO381 protein sequence.	644	99
793	AF121228	Homo sapiens	thyroid hormone receptor-associated protein	1037	100
704	C04022	YIn and an in	complex component TRAP95	1.04	162
794 795	G04072 Y69384	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 8153. Amino acid sequence of a 14274 receptor	124 119	100
		<u></u>	protein.	L	<u> </u>
796	W40215	Homo sapiens	Human macrophage antigen.	1358	99

SEQ	Accession	Species	Description	Smith-	1%
SEQ ID	No.	apecies	Description	Waterman	Identity
NO:	10.	1		Score	Idealthy
797	AF258340	Homo sapiens	hepatocellular carcinoma-associated antigen 112	1151	99
	AF159615			461	98
798		Homo sapiens	FGF receptor activating protein 1	797	99
799	Y59863	Homo sapiens	Human normal uterus tissue derived protein 26.	1	1
800	W70459	Homo sapiens	Human T1-receptor ligand III splice variant 2.	572	92
801	L00073	Homo sapiens	renin	1913	93
802	P92219	Homo sapiens	CRI protein.	11963	97
	L	(human)			
803	X15357	Homo sapiens	ANP-A receptor preprotein (AA -32 to 1029)	5199	98
804	W64473	Homo sapiens	Human secreted protein from clone EC172_1.	4018	95
805	AJ243874	Homo sapiens	oligophrenin-4	2067	100
806	G01731	Homo sapiens	Human secreted protein, SEQ ID NO: 5812.	284	100
807	Z24680	Homo sapiens	garp	1562	83
808	AF171669	Homo sapiens	glycoprotein-associated amino acid transporter	1364	90
	[1 -	LAT2		{
809	W70321	Homo sapiens	Secreted protein CC198_1.	1154	96
810	W74843	Homo sapiens	Human secreted protein encoded by gene 115	855	99
•	{		clone HOVBA03.	1	1
811	AF108831	Homo sapiens	K:Cl cotransporter 3	4561	100
812	AF092135	Homo sapiens	PTD014	862	100
813	AF283772	Homo sapiens	similar to Homo sapiens ribosomal protein L10	784	100
013	A 203772	Tionio sapicus	encoded by GenBank Accession Number	/04	100
	1	[L25899		1
814	G01563	Homo sapiens	Human secreted protein, SEQ ID NO: 5644.	330	100
815	AF051151	Homo sapiens	Toll/interleukin-1 receptor-like protein 3	3850	99
816	W95630			358	100
810	W95630	Homo sapiens	Homo sapiens secreted protein gene clone	338	100
015	001000	 	gn114_1.		100
817	G01082	Homo sapiens	Human secreted protein, SEQ ID NO: 5163.	549	100
818	AF151800	Homo sapiens	CGI-41 protein	1106	95
819	L00352	Homo sapiens	low density lipoprotein receptor	3980	100
820	X04434	Homo sapiens	IGF-I receptor	5832	99
821	G03844	Homo sapiens	Human secreted protein, SEQ ID NO: 7925.	572	100
822	AF212220	Homo sapiens	TERA	396	48
823	Y50125	Homo sapiens	Human glycophosphatidylinositol-anchored	4897	99
	l		protein GPI-122.		
824	AF156778	Homo sapiens	ASB-3 protein	2675	98
825	AF096322	Homo sapiens	neuronal voltage-gated calcium channel gamma-	1105	100
	1		2 subunit	<u> </u>	
826	Y07972	Homo sapiens	Human secreted protein fragment #2 encoded	1540	100
	!	-	from gene 28.	1	ł
827	AB032013	Homo sapiens	potassium channel Kv8.1	2435	95
828	Y13620	Homo sapiens	BCL9	5284	96
829	Y91474	Homo sapiens	Human secreted protein sequence encoded by	541	98
		•	gene 24 SEQ ID NO:147.	ļ	· ·
830	X54232	Homo sapiens	glypican	1625	87
831	X14830	Homo sapiens	acetylcholine receptor beta-subunit preprotein	2540	100
832	Y71262 .	Homo sapiens	Human chondromodulin-like protein, Zchm1.	1002	100
833	G03873	Homo sapiens	Human secreted protein, SEQ ID NO: 7954.	638	96
834	AC003030	Homo sapiens	R29828 1	1389	93
835	Y38422	Homo sapiens	Human secreted protein.	964	87
836	U41557	Caenorhabditi	glycine-rich	85	36
330	031331	s elegans	P-)	"	1 50
837	AL121889	Homo sapiens	dJ1076E17.1 (KIAA0823 protein (continues in	998	75
ω,	W121007	Tronio sabiens	AL023803))	""	1 "
838	AJ011415	Homo sapiens	plexin-B1/SEP receptor	1580	60
839	W80398	Homo sapiens	A secreted protein encoded by clone cw1543 3.		67
				1105	
840	G00862	Homo sapiens	Human secreted protein, SEQ ID NO: 4943.	255	92
841	G02650	Homo sapiens	Human secreted protein, SEQ ID NO: 6731.	644	97
842	AF036717	Homo sapiens	FGFR signalling adaptor SNT-1	2629	99
843	Y73446	Homo sapiens	Human secreted protein clone yc27_1 protein	1089	100
	<u> </u>	<u> </u>	sequence SEQ ID NO:114.	L	L
844	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	357	69
845	AF151810	Homo sapiens	CGI-52 protein	1443	88
046	X83378	Homo sapiens	putative chloride channel	1620	99
846 847	AC004883	Homo sapiens	similar to general transcription factor 2I; similar	655	96

SEQ	Accession	Species	Description	Smith-	%
ID	No.	ļ		Waterman	Identity
NO:		<u> </u>		Score	
			to AF038969 (PID:g2827207)		
848	X99886	Homo sapiens	monocyte chemotactic protein-2	160	76
849	AC005587	Homo sapiens	similar to mouse olfactory receptor 13; similar to	963	98
850	AB038237	Homo sapiens	P34984 (PID:g464305)	 	
851	AF124490	Homo sapiens	G protein-coupled receptor C5L2 ARF GTPase-activating protein GIT1	1767	100
852	Y86217	Homo sapiens	Human secreted protein HWHGU54, SEQ ID	3415 1189	98
652	100217	Homo sapiens	NO:132.	1189	99
853	AF224741	Homo sapiens	chloride channel protein 7	3748	99
854	X17094	Homo sapiens	furin (AA 1-794)	3550	99
855	W78245	Homo sapiens	Fragment of human secreted protein encoded by	1245	99
			gene 19.		
856	R97569	Homo sapiens	Interleukin-2 receptor associated protein p43.	1926	100
857	Y41765	Homo sapiens	Human PRO1083 protein sequence.	3211	99
858	AF057306	Homo sapiens	transmembrane proteolipid	481	84
859	AK025116	Homo sapiens	unnamed protein product	374	69
860	Y41312	Homo sapiens	Human secreted protein encoded by gene 5 clone	824	100
862	Y25776	Nome	HLDRM43.	1005	
863	Y74188	Homo sapiens Homo sapiens	Human secreted protein encoded from gene 66.	895	99
003	1 /4100	Hollio sapiens	Human prostate turnor EST fragment derived protein #375.	96	30
864	AF167473	Homo sapiens	heme-binding protein	870	99
865	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	211	67
866	X54870	Homo sapiens	Type II integral membrane protein	1201	100
867	G00700	Homo sapiens	Human secreted protein, SEQ ID NO: 4781.	640	99
868	Y07894	Homo sapiens	Human secreted protein fragment encoded from	388	88
-			gene 43.		
869	J00123	Homo sapiens	preproenkephalin (1349	95
870	Y91632	Homo sapiens	Human secreted protein sequence encoded by	1048	98
0=1			gene 25 SEQ ID NO:305.		
871	L04311	Homo sapiens	GABA-alpha receptor beta-3 subunit	237	93
872 873	Y29988	Homo sapiens	Human cytokine family member EF-7 protein.	960	94
874	AF161382 G03412	Homo sapiens Homo sapiens	HSPC264	1124	99
875	Y27572	Homo sapiens	Human secreted protein, SEQ ID NO: 7493. Human secreted protein encoded by gene No. 6.	464 573	100 96
876	M15530	Homo sapiens	B-cell growth factor	171	56
877	W63681	Homo sapiens	Human secreted protein 1.	1652	99
878	L27867	Rattus	neurexophilin	1448	98
		norvegicus		1	~
879	Y10835	Homo sapiens	Amino acid sequence of a human secreted	321	100
			protein.		
880	W88991	Homo sapiens	Polypeptide fragment encoded by gene 144.	936	100
881	AF118670	Homo sapiens	orphan G protein-coupled receptor	1971	100
882	AF208865	Homo sapiens	EDRF	528	100
883	Y18462	Homo sapiens	cathepsin L	209	72
884	Y94950	Homo sapiens	Human secreted protein clone dh1073_12 protein	348	100
885	AF070661	Homo sapiens	sequence SEQ ID NO:106. HSPC005	404	100
886	Y04315	Homo sapiens	Human secreted protein encoded by gene 23.	385	100
887	X92744	Homo sapiens	hBD-1	375	100
888	Y22496	Homo sapiens	Human secreted protein sequence clone	994	94
	1	suprous	cn621_8.	//-	~
889	Y41293	Homo sapiens	Human soluble protein ZTMPO-1.	4595	99
890	G03714	Homo sapiens	Human secreted protein, SEQ ID NO: 7795.	147	63
891	AF208856	Homo sapiens	BM-014	1012	99
892	U29195	Homo sapiens	neuronal pentraxin II	2002	98
893	X68149	Homo sapiens	Burkitt lymphoma receptor 1	1953	100
894	Y94914	Homo sapiens	Human secreted protein clone pw337_6 protein	537	100
			sequence SEQ ID NO:34.		
200	W61630	Homo sapiens	Clone HNFGW06 of EGFR receptor family.	326	63
895	104110		CANCILL D manetide management	481	100
896	M24110	Homo sapiens	GOS19-2 peptide precursor		
	M24110 Z68747 AF186112	Homo sapiens Homo sapiens Homo sapiens	imogen 38 neurokinin B-like protein ZNEUROK1	2018	99

SEQ ID NO:	Accession No.	Species	Description	Smith- Waterman Score	% Identity
900	P60657	Homo sapiens	Sequence of human lipocortin.	1835	100
901	M27288	Homo sapiens	oncostatin M	1297	99
902	W85737	Homo sapiens	Polypeptide with transmembrane domain.	749	100
903	G01349	Homo sapiens	Human secreted protein, SEQ ID NO: 5430.	650	99
904	Y00261	Homo sapiens	Human secreted protein encoded by gene 4.	1133	99
905	AF039688	Homo sapiens	antigen NY-CO-3	771	99
906	AB007836	Homo sapiens	Hic-5	2544	100
907	AB017507	Homo sapiens	Apg12	224	100
908	AK000056	Homo sapiens	unnamed protein product	1537	98
909	Y86299	Homo sapiens	Human secreted protein HFOXB55, SEQ ID NO:214.	427	100
910	AF231023	Homo sapiens	protocadherin Flamingo I	7393	99
911	Y14134	Homo sapiens	Vascular endothelial cell growth inhibitor beta protein sequence.	1319	100
912	Z90420	Homo sapiens	Human GDF-3 (hGDF-3) polypeptide encoding cDNA.	1950	100
913	Y19757	Homo sapiens	SEQ ID NO 475 from WO9922243.	1361	100
914	G03172	Homo sapiens	Human secreted protein, SEQ ID NO: 7253.	112	48
915	U14971	Homo sapiens	ribosomal protein S9	886	90
916	AF172854	Homo sapiens	cardiotrophin-like cytokine CLC	1204	99
917	AC005525	Homo sapiens	F22162 1	1963	100
918	AF166350	Homo sapiens	ST7 protein	4711	99
919	Y87285	Homo sapiens	Human signal peptide containing protein HSPP-62 SEQ ID NO:62.	430	100
920	Y36131	Homo sapiens	Human secreted protein #3.	465	88
921	AF193766	Homo sapiens	cytokine-like protein C17	724	100
922	Y95013	Homo sapiens	Human secreted protein vc48_1, SEQ ID NO:66.	357	100
923	X75208	Homo sapiens	protein tyrosine kinase-receptor	5256	100
924	Y96202	Homo sapiens	IkappaB kinase (IKK) binding protein, Y2H56.	813	98
925	AB039886	Homo sapiens	down-regulated in gastric cancer	785	78
926	G03368	Homo sapiens	Human secreted protein, SEQ ID NO: 7449.	55	50
927	Y48606	Homo sapiens	Human breast tumour-associated protein 67.	539	100
928	Y36151	Homo sapiens	Human secreted protein #23.	668	100
929	AF110399	Homo sapiens	elongation factor Ts	1666	100
930	AF210317	Homo sapiens	facilitative glucose transporter family member GLUT9	2763	99
931	Y73328	Homo sapiens	HTRM clone 082843 protein sequence.	931	100
932	G01959	Homo sapiens	Human secreted protein, SEQ ID NO: 6040.	274	100
933	U47924	Homo sapiens	B-cell receptor associated protein	1469	100
934	G03827	Homo sapiens	Human secreted protein, SEQ ID NO: 7908.	529	93
935	AB039371	Homo sapiens	mitochondrial ABC transporter 3	196	63
936	X56385	Canis	rab8	1064	100
,,,	1.2000	familiaris	1250	1.00.	1 200
937	B08906	Homo sapiens	Human secreted protein sequence encoded by gene 16 SEQ ID NO:63.	117	44
938	M13692	Homo sapiens	alpha-1 acid glycoprotein precursor	1064	99
939	Y53886	Homo sapiens	A suppressor of cytokine signalling protein designated HSCOP-6.	515	42
940	Y16630	Homo sapiens	Human Putative Adrenomedullin Receptor (PAR).	1904	99
941	AC005102	Homo sapiens	small inducible cytokine subfamily A member 24	627	99
942	M12886	Homo sapiens	T-cell receptor beta chain	1289	81
943	AF226046	Homo sapiens	GK003	1049	98
944	Y36078	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 463.	667	100
945	M22877	Homo sapiens	cytochrome c	565	100
946	W67869	Homo sapiens	Human secreted protein encoded by gene 63 clone HHGDB72.	551	93
947	W67859	Homo sapiens	Human secreted protein encoded by gene 53 clone HBMCL41.	283	100
948	W85726	Homo sapiens	Novel protein (Clone BG33_7).	789	100
949	AJ242015	Homo sapiens	eMDC II protein	4236	100
950	G04075	Homo sapiens	Human secreted protein, SEQ ID NO: 8156.	567	99

SEQ ID	Accession No.	Species	Description	Smith- Waterman	% Identity
NO:	1 1000			Score	
951 952	AF110645 Y36111	Homo sapiens Homo sapiens	candidate tumor suppressor p33 ING1 homolog Extended human secreted protein sequence, SEQ	1314	70
932	130111	Homo sapiens	ID NO. 496.	402	//
953	AB012109	Homo sapiens	APC10	990	100
954	AF246221	Homo sapiens	transmembrane protein BRJ	1405	100
955	AF054986	Homo sapiens	putative transmembrane GTPase	1883	100
956	W74726	Homo sapiens	Human secreted protein fg949 3.	1879	100
957	Y27096	Homo sapiens	Human viral receptor protein (ACVRP).	1581	100
958	AJ222967	Homo sapiens	cystinosin	1920	100
959	Y53052	Homo sapiens	Human secreted protein clone df202 3 protein sequence SEQ ID NO:110.	587	100
960	G02694	Homo sapiens	Human secreted protein, SEQ ID NO: 6775.	283	100
961	AF151855	Homo sapiens	CGI-97 protein	1214	96
962	U26592	Homo sapiens	diabetes mellitus type I autoantigen	250	65
963	AL050306	Homo sapiens	dJ475B7.2 (novel protein)	3796	100
964 965	AF078859 AB020315	Homo sapiens	PTD004	2089	100
		Homo sapiens	homologue of mouse dkk-1 gene:Acc# AF030433	1466	100
966	X04571	Homo sapiens	precursor polypeptide (AA -22 to 1185)	6580	99
967 968	AF146019 AF071002	Homo sapiens Homo sapiens	hepatocellular carcinoma antigen gene 520 minK-related peptide 1; MiRP1	993	99
969	AB021227	Homo sapiens	membrane-type-5 matrix metalloproteinase	632 3545	100
970	AF180920	Homo sapiens	cyclin L ania-6a	1579	100
971	AF105365	Homo sapiens	K-Cl cotransporter KCC4	5621	99
972	AF083248	Homo sapiens	ribosomal protein L26 homolog	739	100
973	AJ132429	Homo sapiens	hyperpolarization-activated cyclic nucleotide gated cation channel hHCN4	6295	100
974	W61619	Homo sapiens	Clone HTPEF86 of TM4SF superfamily.	454	100
975	AF155100	Homo sapiens	zinc finger protein NY-REN-21 antigen	2261	100
976	AF275948	Homo sapiens	ABCA1	11763	99
977	AB026891	Homo sapiens	cystine/glutamate transporter	2552	100
978	AF117657	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP80	3348	99
979	AF044201	Rattus norvegicus	neural membrane protein 35; NMP35	1570	92
980	AF119297	Homo sapiens	neuroendocrine-specific protein-like protein I	1170	99
981	AF155652	Homo sapiens	potassium channel modulatory factor	1983	99
982	W88499	Homo sapiens	Human stomach carcinoma clone HP10412- encoded protein.	1553	99
983	Z.56281	Homo sapiens	interferon regulatory factor 3	2012	98
984	AB026125	Homo sapiens	ART-4	2160	100
985	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	172	70
986	AB023888	Homo sapiens	b-chemokine receptor CCR4	1895	100
987	W27291	Homo sapiens	Human H1075-1 secreted protein 5' end.	712	100
988	AF153450	Manduca sexta	juvenile hormone esterase binding protein	226	32
989 990	G03697	Homo sapiens	Human secreted protein, SEQ ID NO: 7778.	194	88
	AF204159	Homo sapiens	potassium large conductance calcium-activated channel beta 3a subunit	1486	100
991	G02061	Homo sapiens	Human secreted protein, SEQ ID NO: 6142.	558	99
992	AL031266	Caenorhabditi s elegans	VM106R.1	327	40
993	Y66749	Homo sapiens	Membrane-bound protein PRO1124.	4730	99
994	G01246	Homo sapiens	Human secreted protein, SEQ ID NO: 5327.	141	77
995	AF133845	Homo sapiens	corin	5811	99
996	AF117756	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP150	4999	100
997	W62066	Homo sapiens	Human stem cell antigen 2.	284	93
998	Y87173	Homo sapiens	Human secreted protein sequence SEQ ID NO:212.	725	100
999	Y13379	Homo sapiens	Amino acid sequence of protein PRO263.	1654	99
1000	Y95008	Homo sapiens	Human secreted protein vf3_1, SEQ ID NO:56.	676	47
1001	AF190167	Homo sapiens	membrane associated protein SLP-2	1747	100

SEQ	Accession	Species	Description	Smith-	%
ID NO:	No.		,	Waterman Score	Identity
1002	G01234	Homo sapiens	Human secreted protein, SEQ ID NO: 5315.	398	96
1003	W73420	Homo sapiens	Human secreted protein encoded by Gene No. 24.	2150	100
1004	X12791	Homo sapiens	19kD SRP-protein (AA 1 - 144)	742	100
1005	M23323	Homo sapiens	membrane protein	642	100
1006	X63745	Homo sapiens	KDEL receptor	326	98
1007	Y35997	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 382.	824	99
1008	AB032918	Hylobates moloch	dopamine receptor D4	92	35
1009	Y91680	Homo sapiens	Human secreted protein sequence encoded by gene 81 SEQ ID NO:353.	1372	99
1010	AL136125	Homo sapiens	dJ304B14.1 (novel protein)	825	98
1011	G03733	Homo sapiens	Human secreted protein, SEQ ID NO: 7814.	379	98
1012	Y17531	Homo sapiens	Human secreted protein clone BL205 14 protein.	818	97
1013	G00724	Homo sapiens	Human secreted protein, SEQ ID NO: 4805.	462	100
1014	AF288092	Naegleria gruberi	haem lyase	114	37
1015	AB045292	Homo sapiens	M83 protein	3867	99
1016 ·	X15940	Homo sapiens	ribosomal protein L31 (AA 1-125)	644	100
1017	Y94873	Homo sapiens	Human protein clone HP02632.	1876	100
1018	AL024498	Homo sapiens	dJ417M14.1 (novel protein)	589	100
1019	X83425	Homo sapiens	Lutheran blood group glycoprotein	3054	99
1020	W03516	Homo sapiens	Prostaglandin DP receptor.	1864	100
1021	G03960	Homo sapiens	Human secreted protein, SEQ ID NO: 8041.	398	100
1022	Y91689	Homo sapiens	Human secreted protein sequence encoded by gene 93 SEQ ID NO:362.	768	100
1023	AE000660	Homo sapiens	hADV36S1	573	100
1024	AF132965	Homo sapiens	CGI-31 protein	1550	100
1025	W92380	Homo sapiens	Human TR-interacting protein S103a.	1466	97
1026	R66278	Homo sapiens	Therapeutic polypeptide from glioblastoma cell line.	830	100
1027	X65614	Homo sapiens	S100P calcium-binding protein	476	100
1028	Y41741	Homo sapiens	Human PRO704 protein sequence.	1323	100
1029	AJ001014	Homo sapiens	RAMPI	806	100
1030	W63682	Homo sapiens	Human secreted protein 2.	1354	99
1031	AK023007	Homo sapiens	unnamed protein product	766	100
1032	W97900	Homo sapiens	Human SR-BI class B scavenger.	2672	99
1033	Y82453	Homo sapiens	Human TGC-440 secretory protein SEQ ID NO:1.	639	99
1034	Y73473	Homo sapiens	Human secreted protein clone yd178_1 protein sequence SEQ ID NO:168.	752	93
1035	Y86468	Homo sapiens	Human gene 48-encoded protein fragment, SEQ ID NO:383.	96	90
1036	U09813	Homo sapiens	mitochondrial ATP synthase subunit 9 precursor	698	100
1037	AJ242832	Homo sapiens	calpain	3699	99
1038	X66403	Homo sapiens	acetylcholine receptor epsilon subunit CHRNE	2574	100
1039	AJ242730	Homo sapiens	polyhomeotic 2	1310	100
1040	AF169968	Mus musculus	DNA binding protein DESRT	1453	80
1041	X52563	Bos taurus	permability increasing protein	383	29
1042	G00368	Homo sapiens	Human secreted protein, SEQ ID NO: 4449.	75	50
1043	G02532	Homo sapiens	Human secreted protein, SEQ ID NO: 6613.	60	53
1044	M94582	Homo sapiens	interleukin 8 receptor B	1850	100
1045	AL080239	Homo sapiens	bG256O22.1 (similar to IGFALS (insulin-like growth factor binding protein, acid labile subunit))	1704	50
1046	AF125101	Homo sapiens	HSPC040 protein	580	100
1047	W74809	Homo sapiens	Human secreted protein encoded by gene 81 clone HMWDN32.	176	100
1048	AL022238	Homo sapiens	dJ1042K10.4 (novel protein)	2201	100
1049 ·	W88667	Homo sapiens	Secreted protein encoded by gene 134 clone HAIBP89.	1559	99
		1 - 1		1	

SEQ	Accession	Species	Description	Smith-	1%
ID	No.			Waterman	Identity
NO:	W78324	 	<u> </u>	Score	
1051	_	Homo sapiens	Fragment of human secreted protein encoded by gene 81.	1318	98
1052	Y21851	Homo sapiens	Human signal peptide-contianing protein (SIGP) (clone ID 2328134).	1643	95
1053	AL163815	Arabidopsis thaliana	putative protein	661	62
1054	Y76200	Homo sapiens	Human secreted protein encoded by gene 77.	100	
1055	AJ276567	Homo sapiens	TC10-like Rho GTPase	262 1160	100
1056	Y27620	Homo sapiens	Human secreted protein encoded by gene No. 54.	154	100 96
1057	D14530	Homo sapiens	ribosomal protein	745	100
1058	AF132000	Homo sapiens	TADA1 protein	1132	100
1059	AL031778	Homo sapiens	dJ34B21.1 (novel BZRP (benzodiazapine receptor (peripheral) (MBR, PBR, PBKS, IBP, Isoquinoline-binding protein)) LIKE protein)	920	100
1060	AF227135	Homo sapiens	candidate taste receptor T2R9	134	33
1061	Y27575	Homo sapiens	Human secreted protein encoded by gene No. 9.	1392	100
1062	Z11697	Homo sapiens	HB15	1088	100
1063	AF123757	Homo sapiens	putative transmembrane protein	819	100
1064	AF155135	Homo sapiens	novel retinal pigment epithelial cell protein	2932	99
1065	Y41674	Homo sapiens	Human channel-related molecule HCRM-2.	936	99
1066 1067	AJ250042	Homo sapiens	Rab5 GDP/GTP exchange factor homologue	2575	100
_	Y36087	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 472.	770	85
1068	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	301	100
1069	Y94959	Homo sapiens	Human secreted protein clone mc300_1 protein sequence SEQ ID NO:124.	301	100
1070	W64535	Homo sapiens	Human leukocyte cell clone HP00804 protein.	2014	99
1071	X03145	Homo sapiens	pot. ORF III	148	50
1072	AL031177	Homo sapiens	dJ889M15.3 (novel protein)	821	91
1073	X82200	Homo sapiens	gpStaf50	249	62
1074	G03213	Homo sapiens	Human secreted protein, SEQ ID NO: 7294.	99	47
1075	Y36233	Homo sapiens	Human secreted protein encoded by gene 10.	506	55
1076 1077	G03187	Homo sapiens	Human secreted protein, SEQ ID NO: 7268.	424	98
1077	L25899 Y91447	Homo sapiens	ribosomal protein L10	332	76
	_	Homo sapiens	Human secreted protein sequence encoded by gene 48 SEQ ID NO:168.	898	97
1079	G01862	Homo sapiens	Human secreted protein, SEQ ID NO: 5943.	290	89
1080 1081	AB039723 AB020527	Homo sapiens	WNT receptor frizzled-3	1376	92
1082	L13802	Homo sapiens	Na/PO4 cotransporter homolog	269	100
1083	W75098	Homo sapiens Homo sapiens	ribosmal protein small subunit	499	80
			Human secreted protein encoded by gene 42 clone HSXB125.	143	81
1084	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	83	51
1085	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	88	43
1086	AF090942 G00517	Homo sapiens	PRO0657	124	64
1088	G04091	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 4598.	129	41
1089	AF140631	Homo sapiens	Human secreted protein, SEQ ID NO: 8172. G-protein coupled receptor 14	126	36
1090	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	364 114	82
1091-	S72304	Mus sp.	LMW G-protein	146	32 83
1092	W88708	Homo sapiens	Secreted protein encoded by gene 175 clone	405	100
1093	W85612	Homo sapiens	HEMAM41. Secreted protein clone fh123 5.	4250	07
1094	Y53012	Homo sapiens	Human secreted protein clone pm514_4 protein	4358 1013	97 99
1095	Y92345	Homo sapiens	sequence SEQ ID NO:30. Human cancer associated antigen precursor from	409	100
1096	AF090942	Homo conicas	clone NY-REN-62.	1.0	
1097	L24521	Homo sapiens Homo sapiens	PRO0657 transformation-related protein	147	60
1098	X56932	Homo sapiens	23 kD highly basic protein	166	58
1099	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	490 83	70
1100	Y02693	Homo sapiens	Human secreted protein encoded by gene 44	149	35 59
1100					

SEQ	Accession	Species	Description	Smith-	1%
ID	No.	openies	Description	Waterman	Identity
NO:		1		Score	100.10.5
1101	AF119851	Homo sapiens	PRO1722	183	72
1102	G04086	Homo sapiens	Human secreted protein, SEQ ID NO: 8167.	207	62
1103	G04063	Homo sapiens	Human secreted protein, SEQ ID NO: 8144.	91	52
1104	X74856	Mus	ribosomal protein L28	128	69
		musculus	•	}	1
1105	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	130	62
1106	G03133	Homo sapiens	Human secreted protein, SEQ ID NO: 7214.	122	48
1107	G03040	Homo sapiens	Human secreted protein, SEQ ID NO: 7121.	69	43
1108	AF039942	Homo sapiens	HCF-binding transcription factor Zhangfei	744	99
1109	AF201951	Homo sapiens	high affinity immunoglobulin epsilon receptor beta subunit	738	94
1110	AF111108	Mus musculus	transient receptor potential 2	223	79
1111	AF119900	Homo sapiens	PRO2822	144	59
1112	Y16589	Homo sapiens	A protein that interacts with presentlins.	265	39
1113	G02872	Homo sapiens	Human secreted protein, SEQ ID NO: 6953.	178	67
1114	Y02999	Homo sapiens	Fragment of human secreted protein encoded by gene 121.	164	63
1115	Y30811	Homo sapiens	Human secreted protein encoded from gene 1.	1217	99
1116	X51394	Xenopus laevis	APEG precursor protein	130	40
1117	M27826	Homo sapiens	neutral protease large subunit	442	65
1118	G03371	Homo sapiens	Human secreted protein, SEQ ID NO: 7452.	72	60
1119	G03602	Homo sapiens	Human secreted protein, SEQ ID NO: 7683.	491	97
1120	Y35906	Homo sapiens	Extended human secreted protein sequence, SEQ	244	97
1121	G03714	Homo sapiens	ID NO. 155. Human secreted protein, SEQ ID NO: 7795.	122	
1122	Y00337	Homo sapiens	Human secreted protein, SEQ ID NO: 7793. Human secreted protein encoded by gene 81.	110	90
1123	AF084830	Homo sapiens	two pore domain K+ channel; TASK-2	703	94
1124	AF212862	Homo sapiens	membrane interacting protein of RGS16	442	88
1125	W64469	Homo sapiens	Human secreted protein from clone CW795 2.	191	53
1126	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	154	100
1127	G01361	Homo sapiens	Human secreted protein, SEQ ID NO: 5442.	165	100
1128	Y84320	Homo sapiens	Human cardiovascular system associated protein kinase-1.	815	99
1129	G02105	Homo sapiens	Human secreted protein, SEQ ID NO: 6186.	88	73
1130	Y32923	Homo sapiens	Transmembrane domain containing protein clone HP01512.	700	100
1131	Y29817	Homo sapiens	Human synapse related glycoprotein 2.	260	91
1132	Y91644	Homo sapiens	Human secreted protein sequence encoded by gene 43 SEQ ID NO:317.	525	96
1133	Y91449	Homo sapiens	Human secreted protein sequence encoded by	542	100
		l same ouplois	gene 49.SEQ ID NO:170.		
1134	AB017908 ·	Homo sapiens	4F2 light chain	2399	93
1135	X51760	Homo sapiens	zinc finger protein (583 AA)	312	55
1136	Y99426	Homo sapiens	Human PRO1604 (UNQ785) amino acid sequence SEQ ID NO:308.	917	72
1137	G03790	Homo sapiens	Human secreted protein, SEQ ID NO: 7871.	102	50
1138	AF155106	Homo sapiens	NY-REN-36 antigen	768	91
1139	AL031055	Homo sapiens	dJ28H20.1 (novel protein similar to membrane transport proteins)	117	50
1140	AF011359	Bos taurus	regulator of G-protein signaling 7	138	96
1141	Y70018	Homo sapiens	Human Protease and associated protein-12 (PPRG-12).	623	100
1142	G04091	Homo sapiens	Human secreted protein, SEQ ID NO: 8172.	113	38
1143	AB030235	Canis familiaris	D4 dopamine receptor	89	48
1144	Y94922	Homo sapiens	Human secreted protein clone pv6_1 protein sequence SEQ ID NO:50.	539	88
1145	X99962	Homo sapiens	rab-related GTP-binding protein	398	96
1146	G03807	Homo sapiens	Human secreted protein, SEO ID NO: 7888.	168	79
1147	G03712	Homo sapiens	Human secreted protein, SEQ ID NO: 7793.	512	85
1148	Y28279	Homo sapiens	Human G-protein coupled receptor GRIR-1.	705	76
1149	U13642	Caenorhabditi	exon 5 similar to transmembrane domain of S.	247	36

SEQ	Accession	Species	Description	Smith-	1%
ID	No.			Waterman	Identity
NO:	ļ			Score	<u> </u>
1150	G03438	s elegans	cerevisiae zinc resistance protein	 	
1151	G01003	Homo sapiens Homo sapiens	Human secreted protein, SEQ ID NO: 7519.	117	62
1152	G03798	Homo sapiens	Human secreted protein, SEQ ID NO: 5084. Human secreted protein, SEQ ID NO: 7879.	181,	80
1153	X88799	Oryza sativa	DNA binding protein	95	63
1154	D85245	Homo sapiens	TR3beta		41
1155	R74272	Homo sapiens	Tumour suppressor protein, p53.	155 341	96 87
1156	Y86265	Homo sapiens	Human secreted protein HUSXE77, SEQ ID	99	
	<u> </u>		NO:180.		41
1157	G02577	Homo sapiens	Human secreted protein, SEQ ID NO: 6658.	263	98
1158	AF104334	Homo sapiens	putative organic anion transporter	185	42
1159	G01393	Homo sapiens	Human secreted protein, SEQ ID NO: 5474.	173	57
1160	W75771	Homo sapiens	Human GTP binding protein APD08.	224	81
1161	AF216833	Homo sapiens	M-ABC2 protein	410	83
1162	W67816	Homo sapiens	Human secreted protein encoded by gene 10 clone HCEMU42.	1156	100
1163	AF119851	Homo sapiens	PRO1722	230	70
1164	Y87252	Homo sapiens	Human signal peptide containing protein HSPP- 29 SEQ ID NO:29.	113	31
1165	W64537	Homo sapiens	Human liver cell clone HP01148 protein.	338	82
1166	AF269286	Homo sapiens	HC6	134	64
1167	Y14482	Homo sapiens	Fragment of human secreted protein encoded by gene 17.	149	51
1168	D90789	Escherichia coli	Dipeptide transport system permease protein DppC.	411	90
1169	R63783	Homo sapiens	TG0847 protein.	344	90
1170	Y45274	Homo sapiens	Human secreted protein encoded from gene 18.	478	98
1171	D64154	Homo sapiens	Mr 110,000 antigen	347	96
1172	AB026256	Homo sapiens	organic anion transporter OATP-B	311	67
1173	G00357	Homo sapiens	Human secreted protein, SEQ ID NO: 4438.	60	52
1174	D87717	Homo sapiens	similar to human GTPase-activating protein(A49869)	178	59
1175	M64716	Homo sapiens	ribosomal protein	391	78
1176	R08330	Homo sapiens	Human IL-7 receptor clone H6.	285	67
1177	L06505	Homo sapiens	ribosomal protein L12	242	72
1178	AJ251885	Homo sapiens	organic cation transporter (OCT2)	276	88
1179	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	155	71
1180	G01207	Homo sapiens	Human secreted protein, SEQ ID NO: 5288.	282	90
1181	AF181856	Rattus	tRNA selenocystcine associated protein	249	62
		norvegicus			<u> </u>
1182	AF161524	Homo sapiens	HSPC176	138	90
1183	G03789	Homo sapiens	Human secreted protein, SEQ ID NO: 7870.	282	66
1184	Y02671	Homo sapiens	Human secreted protein encoded by gene 22 clone HMSJW18.	107	71
1185	G03797	Homo sapiens	Human secreted protein, SEQ ID NO: 7878.	88	69
1186	G03564	Homo sapiens	Human secreted protein, SEQ ID NO: 7645.	118	46
1187	AB032905	Hylobates concolor	dopamine receptor D4	96	37
1188	G00956	Homo sapiens	Human secreted protein, SEQ ID NO: 5037.	292	78
1189	G03258	Homo sapiens	Human secreted protein, SEQ ID NO: 7339.	178	79
1190	G03361	Homo sapiens	Human secreted protein, SEQ ID NO: 7442.	324	76
1191	AF117755	Homo sapiens	thyroid hormone receptor-associated protein complex component TRAP230	187	70
1192	Y70455	Homo sapiens	Human membrane channel protein-5 (MECHP-5).	202	67
1193	G03052	Homo sapiens	Human secreted protein, SEQ ID NO: 7133.	99	42
1194	G02607	Homo sapiens	Human secreted protein, SEQ ID NO: 6688.	192	76
1195	W29661	Homo sapiens	Homo sapiens CI542 2 clone secreted protein.	2001	98
1196	Y14104	Homo sapiens	Human GABAB receptor 1d protein sequence.	239	69
1197	X61972	Homo sapiens	macropain subunit iota	149	90
1198	G00534	Homo sapiens	Human secreted protein, SEQ ID NO: 4615.	145	51
1199	Y86260	Homo sapiens	Human secreted protein HELHN47, SEQ ID NO:175.	1089	89
	ł				

SEQ	Accession	Species	Description	Smith-	1%
m `	No.	•	· · · · ·	Waterman	Identity
NO:	1	l		Score	
1201	G00838	Homo sapiens	Human secreted protein, SEQ ID NO: 4919.	404	50
1202	M27826	Homo sapiens	neutral protease large subunit	202	49
1203	Y73424	Homo sapicas	Human secreted protein clone yi4_1 protein sequence SEQ ID NO:70.	265	61
1204	AF264014	Homo sapiens	scavenger receptor cysteine-rich type 1 protein M160 precursor	625	98
1205	Y36203	Homo sapiens	Human secreted protein #75.	219	59
1206	U78111	Gallus gallus	AQ	205	57
1207	AF095448	Homo sapiens	putative G protein-coupled receptor	416	76
1208	AF116715	Homo sapiens	PRO2829	127	75
1209	AF099137	Homo sapiens	MaxiK channel beta 2 subunit	475	95
1210	AF205718	Homo sapiens	hepatocellular carcinoma-related putative tumor suppressor	423	79
1211	Y27868	Homo sapiens	Human secreted protein encoded by gene No. 107.	224	70
1212	G00719	Homo sapiens	Human secreted protein, SEQ ID NO: 4800.	117	44
1213	G01009	Homo sapiens	Human secreted protein, SEQ ID NO: 5090.	351	73
1214	AF090942	Homo sapiens	PRO0657	124	70
1215	Y14427	Homo sapiens	Human secreted protein encoded by gene 17 clone HSIEA14.	99	77
1216	G03905	Homo sapiens	Human secreted protein, SEQ ID NO: 7986.	173	57
1217	Y57897	Homo sapiens	Human transmembrane protein HTMPN-21.	1173	100
1218	J00194	Homo sapiens	hla-dr antigen alpha chain	454	78
1219	Y59709	Homo sapiens	Secreted protein 76-28-3-A12-FL1.	470	92
1220	W81576	Homo sapiens	EBV-induced G-protein coupled receptor (EBI-2) polypeptide.	725	100
1221	W96745	Homo sapiens	High affinity immunoglobulin E receptor-like protein (IGERB).	650	98
1222	Y35911	Homo sapiens	Extended human secreted protein sequence, SEQ ID NO. 160.	135	31
1223	Y00278	Homo sapiens	Human secreted protein encoded by gene 21.	260	95
1224	AF161422	Homo sapiens	HSPC304	568	90
1225	U14970	Homo sapiens	ribosomal protein S5	202	95
1226	G01733	Homo sapiens	Human secreted protein, SEQ ID NO: 5814.	610	100
1227	AF099973	Mus musculus	schlafen2	333	56
1228 1229	G01218 AF217188	Homo sapiens Mus	Human secreted protein, SEQ ID NO: 5299.	155	81
1230	AF176813	musculus	YIP1B	801	63
1231	X98333	Homo sapiens	soluble adenylyl cyclase	275	100
1232	W74955	Homo sapiens	organic cation transporter	1704	100
1232	Y94940	Homo sapiens Homo sapiens	Human secreted protein encoded by gene 77 clone HOEAS24. Human secreted protein clone yi62 1 protein	212	100
1234			sequence SEQ ID NO:86.	526	
1234	U76618	Mus musculus	N-RAP	482	82
1235	AF044924	Homo sapiens	hook2 protein	380	97
1230	G01459 AF000018	Homo sapiens	Human secreted protein, SEQ ID NO: 5540. adapter protein	417	100
1238	W88633	Homo sapiens	Secreted protein encoded by gene 100 clone HE8EU04.	250	90
1239	W29660	Homo sapiens	Homo sapiens CH27 1 clone secreted protein.	697	98
1240	AF004161	Oryctolagus cuniculus	peroxisomal Ca-dependent solute carrier	154	52
1241	Y92710	Homo sapiens	Human membrane-associated protein Zsig24.	709	97
1242	Y95002	Homo sapiens	Human secreted protein vc34_1, SEQ ID NO:44.	908	88
1243	Y44905	Homo sapiens	Human potassium channel molecule ERG-LP2 partial protein.	325	100
1244	AF284422	Homo sapiens	cation-chloride cotransporter-interacting protein	511	97
1245	Y53629	Homo sapiens	A bone marrow secreted protein designated BMS115.	1888	93
1246	AB039371	Homo sapiens	mitochondrial ABC transporter 3	389	97
1247	Y35911	Homo sapiens	Extended human secreted protein sequence, SEQ	168	39

SEQ	Accession	Species	Description	Smith-	1%
ID `	No.	'	7	Waterman	Identity
NO:	- }	ļ	j	Score	Monthly
			ID NO. 160.	-	+
1248	AF072509	Rattus	glutamate receptor interacting protein 2	559	90
	<u> </u>	погуедісия	1	1	
1249	AF247042	Homo sapiens	tandem pore domain potassium channel TRAAK	661	98
1250	B08974	Homo sapiens	Human secreted protein sequence encoded by	1087	97
			gene 27 SEQ ID NO:131.	-	1
1251	L15313	Caenorhabditi	putative	858	59
		s elegans			1
1252	Y29338	Homo sapiens	Human secreted protein clone it217_2 alternate	278	75
1052	17701500	 	reading frame protein.		
1253	W01730	Homo sapiens	Human G-protein receptor HPRAJ70.	211	92
1254	G03074	Homo sapiens	Human secreted protein, SEQ ID NO: 7155.	294	83
1255 1256	G01818 AF286368	Homo sapiens	Human secreted protein, SEQ ID NO: 5899.	253	91
1257	AF286368 AF220264	Homo sapiens	eppin-l	222	54
1257	G02227	Homo sapiens	MOST-1	87	93
1259	Y07970	Homo sapiens	Human secreted protein, SEQ ID NO: 6308.	281	78
1239	10/9/0	Homo sapiens	Human secreted protein fragment #2 encoded	81	94
1260	R95332	Homo sapiens	from gene 26.	000	1
	103332	riomo sapiens	Tumor necrosis factor receptor 1 death domain ligand (clone 3TW).	986	100
1261	AF140674	Homo sapiens	zinc metalloprotease ADAMTS6	155	<u> </u>
1262	U28369	Homo sapiens	semaphorin V	172	36
1263	Y07049	Homo sapiens	Renal cancer associated antigen precursor	237	67
1205	10,015	110illo sapielis	sequence.	288	71
1264	Y36153	Homo sapiens	Human secreted protein #25.	187	80
1265	Y78114	Homo sapiens	Human cytokine signal regulator CKSR-2 SEQ	723	93
		-10	ID NO:2.	123	93
1266	Y13397	Homo sapiens	Amino acid sequence of protein PRO334.	191	100
1267	AF030558	Rattus	phosphatidylinositol 5-phosphate 4-kinase	859	95
		norvegicus	gamma	039	33
1268	U73167	Homo sapiens	candidate tumor suppressor gene LUCA-1	159	96
1269	AF190664	Mus	LMBR2	552	76
	1 .	musculus		552	10
1270	AL050332	Homo sapiens	dJ570F3.1 (homolog of the rat synaptic ras	820	98
			GTPase-activating protein p135 SynGAP)		"
1271	G02126	Homo sapiens	Human secreted protein, SEQ ID NO: 6207.	131	95
1272	AF125533	Homo sapiens	NADH-cytochrome b5 reductase isoform .	253	92
1273	AL035661	Homo sapiens	dJ568C11.3 (novel AMP-binding enzyme	1280	100
	1		similar to acetyl-coenzyme A synthethase		ļ
1071	1 100000000	ļ	(acetate-coA ligase))		}
1274	AF064748	Mus	S3-12	3523	61
1275	Dizect	musculus			
1276	D17554	Homo sapiens	TAXREB107	377	78
12/0	Y30715	Homo sapiens	Amino acid sequence of a human secreted	643	90
1277	AF146760	Homo sapiens	protein.		
1278	Y05069		septin 2-like cell division control protein	707	100
1279	X59668	Homo sapiens Oryctolagus	Human PIGR-2 protein sequence.	281	46
-2.,	12,000	cuniculus	aorta CNG channel (rACNG)	267	85
1280	G01051	Homo sapiens	Human secreted protein, SEQ ID NO: 5132.	400	00
1281	G03411	Homo sapiens	Human secreted protein, SEQ ID NO: 5132. Human secreted protein, SEQ ID NO: 7492.	489 120	98 43
1282	AF055084	Homo sapiens	very large G-protein coupled receptor-1		
1283	AF117814	Mus	odd-skipped related 1 protein	1635 357	100
	{	musculus	and arribbon totation I biolem	331	98
1284	U87318	Xenopus	NaDC-2	535	60
		laevis		دود	00
1285	AF061346	Mus	Edp1 protein	452	68
		musculus	• • • • • • • • • • • • • • • • • • • •	.52	J.
1286	AB030182	Mus	contains transmembrane (TM) region	582	68
	<u></u> .	musculus	() • • • • • • • • • • • • • • • • •		55
1287	A13595	synthetic	immunosuppresive protein PP15	185	97
		construct			
1288	AF254411	Homo sapiens	ser/arg-rich pre-mRNA splicing factor SR-A1	837	100
1289	AF084205	Rattus	serine/threonine protein kinase TAO1	319	98
	1	norvegicus	-		

SEQ	Accession	Species	Description	Smith-	% Idania
ID D	No.	[Waterman Score	Identity
NO:		<u> </u>	The state of the s	523	100
290	AF038563	Homo sapiens	membrane associated guanylate kinase 2 double-stranded RNA specific adenosine	468	100
1291	AF034837	Homo sapiens	deaminase		
292	M15888	Bos taurus	endozepine-related protein precursor	937 636	87
293	AB010692	Arabidopsis thaliana	ATP-dependent RNA helicase-like protein		
1294	AF209923	Homo sapiens	orphan G-protein coupled receptor	1570	100
295	W67828	Homo sapiens	Human secreted protein encoded by gene 22 clone HFEAF41.	504	98
296	AC004832	Homo sapiens	similar to 45 kDa secretory protein; similar to CAA10644.1 (PID:g4164418)	648	65
1297	X80035	Oryctolagus cuniculus	cysteine rich hair keratin associated protein	575	70
298	G02645	Homo sapiens	Human secreted protein, SEQ ID NO: 6726.	223	97
1299	Y59440	Homo sapiens	Human delta3 fragment #4.	122	32
1300	W70504	Homo sapiens	Leukocyte seven times membrane-penetrating type receptor protein JEG18.	459	81
1301	Y67315	Homo sapiens	Human secreted protein BL89_13 amino acid sequence.	3916	99
1302	M77693	Homo sapiens	spermidine/spermine N1-acetyltransferase	174	96
1303	G01331	Homo sapiens	Human secreted protein, SEQ ID NO: 5412.	254	69
1304	G01491	Homo sapiens	Human secreted protein, SEQ ID NO: 5572.	747	99
305	AF148509	Homo sapiens	alpha 1,2-mannosidase	602	98
306	G01658	Homo sapiens	Human secreted protein, SEQ ID NO: 5739.	333	98
1307	Y90899	Homo sapiens	D1-like dopamine receptor activity modifying protein SEQ ID NO:1.	332	98
1308	AF033120	Homo sapiens	p53 regulated PA26-T2 nuclear protein	348	52
309	Y73388	Homo sapiens	HTRM clone 3376404 protein sequence.	147	66
1310	AF063243	Bos taurus	ribosomal protein L30	296	90
1311	AF224494	Mus musculus	arsenite inducible RNA associated protein	688	70
1312	Y73342	Homo sapiens	HTRM clone 2709055 protein sequence.	1154	100
1313	Y99419	Homo sapiens	Human PRO1780 (UNQ842) amino acid sequence SEQ ID NO:282.	1145	78
1314	AF116667	Homo sapiens	PRO1777	433	j 97
1315	W75100	Homo sapiens	Human secreted protein encoded by gene 44 clone HE8CJ26.	807	97
1316	AJ272078	Homo sapiens	APOBEC-1 stimulating protein	789	100
1317	AB041533	Homo sapiens	sperm antigen	2607	98
1318	U19617	Mus musculus	Elf-1	806	92
1319	U82598	Escherichia coli	ferric enterobactin transport protein	768	100
1320	D90892	Escherichia coli	SORBITOL-6-PHOSPHATE 2- DEHYDROGENASE (EC 1.1.1.140) (GLUCITOL-6- PHOSPHATE DEHYDROGENASE) (KETOSEPHOSPHATE REDUCTASE).	709	100
1321	W67847	Homo sapiens	Human secreted protein encoded by gene 41 clone HPBCJ74.	601	92
1322	AJ276101	Homo sapiens	GPRC5B protein	466	93
1323	AJ276101	Homo sapiens	GPRC5B protein	504	97
1324	Y58628	Homo sapiens		1584	100 89
1325	U91561	Rattus norvegicus	pyridoxine 5'-phosphate oxidase	<u> </u>	
1326	AF125533	Homo sapiens		1606	100
1327	¥32206	Homo sapiens	Incyte clone 2825826.	1531	90
1328	AF151048	Homo sapiens		657	85
1329	Y10530	Homo sapiens	olfactory receptor	1645	100
1330	AF180681	Homo sapiens	guanine nucleotide exchange factor	4314	99
1331	AF111856	Homo sapiens	NaPi-3b	3591	99
1332	Y13583	Homo sapiens	G-protein coupled receptor	2171	100
	AF078866	Homo sapiens		1395	100

SEQ	Accession	Species	Description	Smith-	%
ID	No.	[[Waterman	Identity
NO:				Score	
1334	Y25755	Homo sapiens	Human secreted protein encoded from gene 45.	1380	96
1335	AF152325	Homo sapiens	protocadherin gamma A5	4742	99
1336	X74070	Homo sapiens	transcription factor BTF3	639	81
1337	AF095927	Rattus norvegicus	protein phosphatase 2C	1931	95
1338	G03877	Homo sapiens	Human secreted protein, SEQ ID NO: 7958.	621	100
1339	AL008582	Homo sapiens	bK223H9.2 (ortholog of A. thaliana F23F1.8)	626	100
1340	X61615	Homo sapiens	leukemia inhibitory factor receptor	5820	99
1341	Y01519	Homo sapiens	A carcinogenesis-inhibiting protein.	7528	97
1342	AF207600	Homo sapiens	ethanolamine kinase	2372	100
1343	U54807	Rattus norvegicus	GTP-binding protein	1167	97
1344	AC020579	Arabidopsis thaliana	putative phosphoribosylformylglycinamidine synthase; 25509-29950	3283	51
1345	Y28576	Homo sapiens	Secreted peptide clone pe503 1.	944	100
1346	W74787	Homo sapiens	Human secreted protein encoded by gene 58 clone HHFHN61.	1171	100
1347	M55542	Homo sapiens	guanylate binding protein isoform I	2636	87
1348	AF183428	Homo sapiens	28.4 kDa protein	1329	100
1349	U70669	Homo sapiens	Fas-ligand associated factor 3	167	24
1350	AF295530	Homo sapiens	cardiac voltage gated potassium channel modulatory subunit	562	99

TABLE 3

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Thrconine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1	1351	A	2	337	1	TPSLIHQAPTPCPAGLWG/PPNGHYHGS*PGC HWPQAPHRA***GLLPPRWLGHGLPGGPAAP WAASQWVDGVAGRLPGPAWSWHASGAAPA QPGPL*LLVPGSSGLPDPRDP
2	1352	A	27	100	366	IRNSSIRPMKERETKLSAKHMITCSASYDIRGL QIETT\YHHTPIRMAKIQKT/GHHQC**ECGAT GTLIHGWWGCKVVEPLGKTVWQIPK
3	1353	A	40	3	314	·HASAHASVVLKDNSELEQQLGATGAYRARA LELEAEVAEMRQMLQLEHPFVNGADKLRPD SMYVHLNEL*QSLVENMLLTVVDTH/RTPI*R SCNYTLALILFL
4	1354	A	74	2	292	TASALFSCPDGGSLAGFAGRRASFHLECLKR QKDRGGDISQKTVLPLHLVHHQVAHTFGQAT VTCQQARQSPG*RTNPE/ALQWVLPVSDGWH VLPLP
5	1355	A	78	114	850	ENCRVASNLPGVFFSEDTAQSGSYMRISAHPP NAGGEVSNGPKRKLTLMLNFSLPSSGLNAGA FYALSTLLNRMVIWHYPGEEVNAGRIGLTIVI AGMLGAVISGIWLDRSKTYKETTLVVYIMDT GGAWWCYTFYLGTGDTCG*CFITAGTMGFF MTGYLPLGFEFAVEL\SYPESEGISSGLLNISA QVFGIIFTISQGQIIDNYGTKPGNIFLCVFLTLG AALTAFIKADLRRQKANKETLEN
6	1356	A	81	97	376	EWFSYMLGSNMSVYHSP*SLEPLCKVLSES*A YLRVPFIRILLNAR*IRKAYKRMSLEIKLLIIRE *CLFQEMGLSLQWLYSARGDFFRATSRL
7	1357	A	93	2	872	TLSSACLIGDAWKELTIVAGAVSNOLLVWYP ATALADNKPVAPDRRISGHVGIIFSMSYLESK GLLATASEDRSVRIWKGGDLRVPGGRVQNIG HCFGHSARVWQVKLLENYLISAGEDCVCLV

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion WSHEGEIL QAFRGHQGRGIRAIAAHERQAWV ITGGDDSGIRL WHLVGRGYRGLG/DLGSLLO
						VP**ARYTQGCDSGWLLATAGSD*YRGPVSL *RRGQVLGAAARG*TFPVLLPAGGSSWSRGL RIVCYGQWGRSCQGCPHQHSNCCCGPDPVS WEGAQLELGPAWL
8	1358	A	106	3	350	FSSLLSGRISTLRDETGAILIDGDPAACAPIIKF LLTEELHLRGVSIYVLRHEAQIYGITPL\VCAL L!/CRRL*SDSCMRAALNDRGLYQVLILDGLV QCLGFVDSDSRKMVSTLT
9	1359	A	115	49	186	QAWAIFKGKYKEGDTGGPAVWKTRLRCALN KSSEFNEGPERERMDV
10	1360	A	123	2	1249	KGCRTQEKVDRTEVIRTCINPVYSKLFTVDFY FEEVQRLRFEVHDISSNHNGLKEADFLGGME CTLGQIVSQRKLSKSLLKHGNTAGKSSITVIA EELSGNDDYVELAFNARKLDDKDFFSKSDFF LEIFRMNDDATQQLVHRTEVVMNNLSPAWK SFKVSVNSLCSGDPDRRLKCIVWDWDSNGK HDFIGEFTSTFKEMRGAMEGKQVQWECINPK YKAKKKNYKNSGTVILNLCKIHKMHSFLDYI MGGCQIQFTVAIDFTASNGDPRNSCSLHYIHP YQPNEYLKALVAVGEICQDYDSDKMFPAFGF GARIPPEYTDSHDFAINFNEDNPECAGIQGVV EAYQSCFPKAPTFTGPTNICPHSSRKVAKFRR SEGN*HQGRAFAIIFILVDPGQVGVYSQDMGP DNPGGHFV
11	1361	A	147	614,	9	ACARKQLLGRTVFIWFVGQLLGGELKGYSKT NTTSSRPASSRG\TLSSSSSSSSSLTKDALPSSL KSDSTTITSGLVFPFRSLCVNPAKSSVSESVSSI KILLSSSVKYLE*KRTSCCFPDSSESKLSQLSS DERVSMGTSSRKPTNSSSSLGALKMSATS*G SGSESPTPFFLTGLQSPPSTRPREPGLTTARNS TTLTRDC
12	1362	A	177	12	416	LIPSEPALDSLVDPRVRSRKQPFVIYPVYDTAI DTKIHFSLLDGNVGEPDMSAGFCPNHKAAM VLFLDRVYGIEVQDFLLHLLEGGFLPDLRAA ASLDT/AEIGAMDFLLS*LFTLCLMMFFFIYPFI NLLTMNVY
13	1363	A	249	535	105	WTFHRHLSPAPLIVCDQGTCVVSYYPQNIVQ MPDTQMEQGLN/HLFLDGNA*PHSVECYCPS TFEIAIKITSFVLYFHRYRAPEVLLRSSVYSSPI DVWAVGSIMAELYMLRPLFPGTSEVDEIFKIC QVLGTPKKVSTLVPKLL
14	1364	A	254	572	201	YLLTXIGNLMMLLVINADSCLRTXM*FFLGH FFFLDICYSSVTAQDAAEFPVS*KPILVWGYIT *SFFFIFSWGTNGCLLSAITYACYAAICHPLLS TMVMNRPLCTATVNATNKMGFLNSQVN
15	1365	A	257	425	68	THAKFLNKKFNIPKLVILPKLVYIVKAIPTKM AIEFILIECDQNIT/KLICENT*KNIAKNI*KRRV TFTPIET*HPVKQMIKWQ*LTAWLRNRGYKKI KQTPNSETAPSVCRNLVFDKCG
16	1366	A	263	104	481	FCIFRTTEEDRGGDDCVVSVWTKQRNNSCVK SKDVFSKPVNIFWALEESVLGVKARQPKPFFA AGNTFEMTCKVSSKNIKSPRYSVLIMAEKPV GDLSSPNETKYIISLDQDSVVKLENWTDASRV
17	1367	A	298	68	208	RKRTNNPIKLDKKFEHFKNEDI*ITSKHTKMW VSSLAMKEMLTKTTM
18	1368	A	300	904	ī	LVVGITGTRHHARVIFIFLVETGFPHVGQAGL ELLTSGDPPALASQSAGITGMSHCARPKGHFG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cystcinc, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion IHLK*MFYTMSQKMP*PTINLILLLIPGNLNIF KPNMGWLGPKTAFV*KDEVLSGIPFAKGRCR WK*DY*C/LQEVTDPIMEKGKKKKRTASFFK GQPHQSTNALLRCVR*RYHLS\TVETAGLP*KNTGHIPGQPFLFKLVFKC*NVICI**QYKW*Q NIGVKNKSFCPH*SSSPSL*FIGHHSRNF/CSFK TEPHSVVQAGGQWRNLSSLQAPPPGLMPLSR
19	1369	A	302	3	445	ISLMSSWDYRRPPQ NSPSRWAKIQMFEHTFCG*GCG/ER/NVHIHCS WICRLRPLLWRAVREYLSKLKNAELSFDPGV SLLRIYAIDMPTSI*DEKEALLFAFLAFHE*HC KSRIWAVIQ/CIHLWDWLRKL*CFHRMKFYA AV*NKPRHLLSHIWKDVQNILLK
20	1370	A	304	1	1339	FFFCGKEVPLFEQNKHPGPRATTSPGA/HARA LLSAGEFTAGVGLSP*AIHSFVWLCTFIQHGA GGPCHQPGGSPGPWMHTTQAGHLWEGAYPG GSSTWHQVPGQLGGSWGPRERSLLGSFIKCSP CPHPPGFRLWMSPNQKPPTENPGVMGRVWR LMPGESPLIWEAEGKEDHLSPEGQGHSE/PVA PLHSSLGNTVKP*PKNQKPKQNRSRHGQ\GF MAGQGQSRPAAR*PPCPALTPASHSAGTWPP RICRTVPGGPCPSPSGFRSCRR*GFSA*TRSWP DAEPPSTPDTAPRCCTQSDTSSQGPQ*S*WRR CRALPGRLCSAPAAGLRRARPRLSESRRGNSP PASPAAASARCPSWGPSCPARPPSRPAAGTEP AAPSRCTAWLRGEREPGPRPPGRRPRSGRGP VSFAPEVLSLPAVRQTKSWRWRNEEEITRPW ALVRSRGG
21		A	326	799	1587	GSQVLPPPPSQDSATLPQDA*GPRAAPGQPVC E*GLQGAGVRRLRGEVLCQPQP*GAL*EQCLP HLSFSPRQGAAPDTEPSAWGPAPTGATGPGLP LRHVRLFSAGAPRGAATPCPPALLHGPAWPP ARPMFRGHPPVRPLGPWGKVAAGPRALCLA GVPAVQGECATKPSG*GL*PAHLRGPPGPEVL QWHWQLSAGRDPVPAEDPPL*EGPLGPGGPA AAQAEPGADPEPEDKDQAAESRPAGAMSLSA QGSGPVGGQGLR
22	1372	A	327	146	652	PHLENPHPEHSFPGAPLT*STLSWSILSPREPSP GAPCYPGHPHLENPHLEHLLTWRTVTWSTLL PGAPCYPEHPHLEHPLTWSTPHLEHPSPGEPL SCRTPTRSILHRDHPLP*CLSTEESPI*GWGSLP APPSTPLVLDVAPPGPQPASSCPGRDSCYSVP GTVVSP
23	1373	A	348	397	2	CIVSSCQGTRKPCHLEDANKINKQSPTLEKIES LQESL*VKQ*LIVAEKYVQILHPRKKYFQRPL NNEKRKMKKRKEEKKKCRERMQRRSKWRR EEKKE*RREE\EERKKEKEDRKERRKETSPRG SRRLLRD
24	1374	A	362	170	352	GRALDTAAGSPVQTAHGLPSDALAPLDDSMP WEGRTTAQWSLHRKRHLARTLLVSRVRGPQ
25	1375	A		373	128	YLITTILETGYLWKNRHSDQ*KRTENPERDQH KYPKVDFCKSNSMKNRLCNKWHWTNWIFTD KKINLNLKPHTKLTPNIKKN
26	1376	A	397	383	165	EVKNTNPFIFSGTNLTIWIRSI*RKSDEINQRTK *MEKYSISLDRRLNTVKMSFLPNLIYKFNTISI KIPANF
27	1377	A	406	103	380	KSKATGYMVNI*KLIV\FLYANDEQLEIEMNK IVP\FNGSKNKIAFTNLTKYQNIQNRHAENYKI LVNKIEDLNKWRNYLLSWIGRRNIINTMT

		1	
 	 		· · · · · · · · · · · · · · · · · · ·

-

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide	100	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	J	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ì	ł	914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine.
[{	ĺ	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	İ	ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
(1	1	1	peptide		/=possible nucleotide deletion, \=possible
28	1378	A	408	sequence	427	nucleotide insertion
20	13/6	^	408	14	427	TICTNKFNNLDEIK/FLERHKLSKLTQEEVENL ITLKTSRETELVINK*VIPHKEKPGPDSFTGEF
	[1	ĺ	[ĺ	YQTFKEEL/I/ILHKLFQTIKYGRILPNSVYETSI
	İ			j	İ	TLKPKPEKDL/KENYRPLPLSNIDAK/LNKTLA
	1			{	ļ	NRI**HIR
29	1379	A	434	395	128	IYSKMCMERQRLNN*ILKKNKVRGIAVPDVK
		l	Į	1	}	VYYKPTVIK/TSWIL*KDSHIVEWNRLENLEID
		L	İ	ł		PN/IKRLILDKGAEATEWRKDSFFRQWQ
30	1380	A	455	2	228	FFFETESHSVTQAGVQWCNPGFKRFSCFGLSS
		1))	SWDYRYAPPRP\ANF*FLVETGFYYVAQAGL
						KLLSPGDLPALAS
31	1381	A	462	393	2	QLMFDKGVKNIH\WGWTPPFTK*YWKNWISI
		1				CRRMNLNPYLSRYIKINSR\KDLTVRPEPIKLV
		Ì				EENTGKTIQDTGLGK*FIAKTSKAQSTKTNK*
		ļ				KRQTRYIKLK\KKSTASKENNRVKRQPLE*EK
32	1382	A	474	125	471	IFAN
32	1302	^	1 4/4	12.0	4/1	VKPYEIAVFLVKPIEYK*HLLSDPAIPLSGI*LK EIKAYT/RRICTPMFAAPVSVIA/RN*KOSK/CO
				ı		KQ*YVHRMEYYTTIKRSEILICTTTWVDFRNT
		1				ILRETDRIHKTTYDVISLI
33	1383	A	488	1825	2	KSACSFICSEEQPASPSPLKPGTYASET\RPRDP
					_	HAAGPRRDSSEAETRRPRGA/DGSGTVVKGT
		1	i .			PGSPAPPCSWGHGG\ETEGAG*CPAAPGTDLR
						APGGSAGS*\GLPSAGGSRGRKGWRAAGROP
						STR*GRPGRHGGRGE*AGHPEPRQSALQSAG
						L/ASSPEPMGAALAEDGSGDSRGAGPRPQE*P
]			PSVLSRS\GS*G*G*AASGTASSPRSHSSRLGPP
						SAGFHGLRCGQPPFAAAPPGPWPGTGRPAGG
						AGSPPAAAGTAPPATRGAQSRRQNRTAGRNA
			[[SPQTAAGAGSPVQWALSRATG*TGETGSWC
			1			AGGTHQATHLTAAWVCPPTWSVRPGGSGPA AGLGR*GRHPAQSPPLPVPRG*PAWPQEAPSP
ļ	•]			SPASSEVALSSGSCWPDQAPGPARGSPPAPLA
						PAWPAAGRGRQR*GRQSAHPPPRR*STAVSL
						SGTS*WRRSP*AGTRTQQC*SPWLVPACSSRP
			[L*RGTRRPSTQQSPQTTGTPGRSAGPGHPRS*
ļ						GGRSPAGTGHLGAQTVASPH*GHWPTALSCL
ĺ			1 1	l	'	WASASPPGPEAPPQTGACIGTNCRYRAASAR
						RSSVAPACA*GWQ*AGSPPAVLRGPP*RVRER
34	1204	Α	407	423		GALTHRPRAPDE
77	1384	A	497	422	2	APGASVGRAQAAEG*RGGPTGRPPSALGVS/E
						AGRAGRAGEGRPVPPAYPLCKSAQTSGPPKA
						RLS\PPLASCGGRGPPGGAACATCAPPAGPAR SSRCRRRSPPE*GPR*PSRPARPSPGSAASRRO
		_		l		KLTPCRCQFRGLCA
35	1385	A	509	156	475	PTPYPGE*OAAFLLRGPGLRPPA/DPSLR/HRN
						LTELVVAVTDENIVGLFAALLAERRVLLTAS
ļ]]	j	ļ	KLSTLTSCDHAFCALLYPMRWEHVLIPTLPPH
				.1		LLDYC*CPPLPRT
36	1386	A	512	3	1631	FFFSFVCHLYCVSPTPGPHGRLATWL/PGLLA
ſ	J					FLGLAAGGQTLCPAGELPGHARAQASGAPGS
ļ			ĺĺ	ſ	ſ	VLIAVPGRRRVIITCGPGPAAPSTRGECPPPAL
						GHTRPARPRPV\PFAPAVPQEPGGQGHGAA/P
1	ľ		İ	1		PATGHSAPRGCPPARAAPTGSATPAPPPAACA
	Į]	ļ	AFHSAWSVPPAGRQQG*RVPAPAFRRTTPGT
	İ				ł	PGQHLLDRPGAPPAQGSGPAPAPPPRLAGPA
	.			1		GPAAPPPGPPAASWHSSLSKSSSSL\GWSPPLP
ł			}	})	VGPGSLQ*TPPPQGPHLSGSCGGTSSWRGQR
			<u> </u>			AAVARRLRSWNACGLSRVAGRSSASYPGRE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid	Predicted end nucleotide location corresponding to last amino acid residue of peptide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
)		residue of peptide sequence	sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						GRPSQSQ*PAGPPGMRGCCLRGW*PSSSGSD GPGPHPASTWLRAGKTGPSPPACGCA*LPPPS VSAAPQSPRTRCPRGCAAAAGLCVLAAAGAS HGA\GLPGVRVHTQRVHIH*GAG/GCQTPRPR LRSLPVLGLPAPRCPVSAHPWHRRSGSSCHA ARLVPRHPAPGCP**TG*\PLITGFPEP*A*GLP NHQAVGLEASGALQAGHRDELPTMVQLLDH
37	1387	Ā	620	828	1	SPDYPLKGRPHAP FRLPLAAGA/RGAAEPRVAVSMAPDPSAKIH WEASPEMQSKCHQKGKNNQTECFNHVRFLQ RLNSTHLYACGTHAFQPLCAAIDAEAFTLPTS
						FEEGKEKCPYDPARGFTGLIIDGGLYTATRYE FRSIPDIRRSRHPHSLRTEETPMHWLNG*EDE AQDDGG*GTISSFLLPWPADHPTPKSPGEPVH SIPVCCQVRGQPQSGGKESPACLKSLSNCLTH VDAEFVFSVLVRESKASAVGDDKVYYFFTE RATEKESGSFTQSRSSHRVARGIPPL
38	1388	A	739	1	427	FRAMVSSTLKLGISILNGGNAEVQ/QGNRGKG TSEEGKEG*EVPV*LPVSPPLPRPLQKMLDYL KDKKEVGFFQSIQALMQTC\GEKVMADDEFT QDLFRFLQLLCEGHNNDFQNYLRTQTGNTTT INIIICTVDYLLRLOESI
39	1389	A	767		1030	TLDLTGPLLLGGVPNVPKDFRGRNROFGGCM RNLSVDGKNVDMAGFIANNGTREGCAARRN FCDGRRRQNGGTCVNRWNMYLCECPLRFGG KNCEQGEWPASSIPPVTAAWEALLLDVPGTT VRGLHIQVRQPLVVYAAFTVDSHRPLQETVL RRAPAPASGVPSPSGVGWDR*AGPAEPSPSTP ATVIISVPWYLGLMFRTRKEDSVLMEATSGG PTSFRLQVTGAPCHQGTC*VGARGRDPMLSG LRVTDGEWHHLLIELKNVKEDSEMKHLVTM TLDYGMDQVSWHLHLLWG*TLPPAQGKTGA SEDKVSVRRGFRGCMQVRGGCGGRGEACPS QAAPRL
40	1390	A	801	69	399	IHKIIIHKEDLNKWKYILCSGMERLSTVMIPVV PQIIYKFNA*Q\VILKFTW*E*GAKITILRKNKL RGLVLVPLSTC*VKYLLDKVLPHIKTYYEAR VNKSVVLVQVTIM
41	1391	A	835	7	195	SMLKERKVFQFPSCLFFQYITWLGPPYHVLFD SSVTNFSIGAK*DILOSVMNCLYAKRIPCVT
42	1392	A	841	1	415	GSTHASGYDKTPDFILQVPVAVEGHIHWIES KASFGDECSHHAYLHDQFWSYWNSLKHRTW QGIGTVASNLSQL*TLNAPFPELLLFRSLARTG FVLT*\RFGPGLVIYWYGFIQELDCNRERGILL KACFPTNIVTL
43	1393	A	845	358	92	PALSPAPVPQKKGSPLPLDPCLGPSSWLLSVG LGWPRL*PRRGPGDPGSLPATPPLLTPPHTLLP QRPMLPPSHAGLARPPPPEPISVP
44	1394	A	853	452	1	LPQYCFFPRLSPKSKLVKHSAL**PSALKPPTK SPRCIPRTSLYFTICC/PPALQL/SPIEDPPAIYRS PPTHMLRSASQPLNQAPTLVKGHPPSRFLQG QVSCPPQPTLPREKPLPLHLRPPPRPAQPPLPR PLTFSTRRNVDPEIPERFR
45	1395	A	894	379	162	GVYPPTVFDNYSVQTSVDGQIVSLNTWDTAG QEEYD/RLRTLS*PQTSIFVICFSIGNLEFPIYGT WLSMSMGK
46	1396	A	900	1	366	TTKKTLISNNVSSRSLPILPELKAFSLAFNDPL EIQKYMRT/DQ*CVTHDISLYIVTKLALIFLIPR VFLFHQLNIT**CLHFFTMTTFIAIPFSFLFLGR

SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ł				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	ļ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			[.	peptide	1	/=possible nucleotide deletion, \=possible
	Į			sequence		nucleotide insertion
						D/KSLAMLPRLVSNSWPQVILPP
47	1397	Α	944	162	2	QLQNLASRGCL*SQLLRRLRRENRLNPGGGG
	ļ	ļ	1	ĺ	f	CSEIAP\CTPAWVTQRDFFRKKK
48	1398	A	963	216	308	HFTPDRIAIVKNTRDSHCWRGC*EEGAPARC
49	1399	A	967	466	1	PRKRESWWGERLP/PRGFPPAAEDAPAPGWK
	[į			[•	GRKHASRTARAHVFHPIRQSIRSPVRGRPGDP
ļ	Ì			`		RAAHTRSAGTRLQCKASRGG*GKGPAPTR*E
1		ļ			ļ	GGPGSAPAPLPASSGCSLFPDSSPWTPPPPAPG
L						AAAAQP**TPRCPAALRAGAHIGRVGRPY
50	1400	Α	973	45	421	EKCIQALDVFVFCYIDHSSHCLMSCD*E/DQA
	İ	İ	ì			LNFMPLEMEPKMSKLAFGCQRSSTSDDDSGC
					İ	ALEEYAWVPPGLRPEQIQLYFACLPEEKVPY
	,		İ		<u> </u>	VNSPGEKHRIKQLLYQLPPHDNEVRYCQSLSE
	<u> </u>				<u> </u>	E
51	1401	Α	992	2095	194	IRIRHEAARSCLGCAAGHVPAPGLRLLPTVRG
						PPGRRGPAAPGCVCY*SGESTFVSHVPQRMA
		ſ				WPGSAPPRGFHPLQSQTSPSDTVSSPQLSKEE
		l				DGPGWEHPLSSSL*SLGQAGGNH*QPEELAG
						WEPRGPPSLAPSSPT/TMWTALVLIWIFSLSLS
						ESHAASNDPRNFVPNKMWKGLVKRNASVET
i		i	1			VDNKTSEDVTMAAASPVTLTKGTSAAHLNS
1		1				MEVTTEDTSRTDVSEPATSGVAADGVTSIAPT
						AVASSTTAASITTAASSMTVASSAPTTAASST
		1				TVASIAPTTAASSMTAASSTPMTLALPAPTST
				•		STGRTPSTTATGHPSLSTALAQVPKSSALPRT
}	}	ŀ)			ATLATLATRAQTVATTANTSSPMSTRPSPSKH
	1					MPSDTAASPVPPMRPQAQGPISQVSVDQPVV
		1				NTTNKSTPMPSNTTPEPAPTPTVVTTTKAQAR
ł	ł	l	,	•	!	EPTASPVPVPHTSPIPEMEAMSPTTQPSPMPYT
						QRAAGPGTSQAPEQVETEATPGTDSTGPTPRS
ł						SGGTKMPATDSCQPSTQGQYMV/DHH*APHP
1	l	İ	ł			GRGRQNSPSGGAVTRGDPFHHSLGFVCPAGL
						*ELQEEGLHPGGLLNQRDVCGLRNVRGAGA WREAWPLPRPFLLPLRPNOVLPNSFGAIEEIC
•						OMLKHI
52	1402	A -	994	1	462	ESGEFLVSFTLKKPTNVFHHINGMKFFNK/LIF
1 32	1702	^	""	•	702	*SHTDIAFYKIQHPFMLKALTKWA*EGT*PDR
1		!				RYLH*SLRLNGEQLKTFPLRSGMR*G/CAILPL
		ļ				VLNAMLSIVPAVVPAGKTRHEKEITCPLIGGE
		ĺ	:			EK*FS*FVGDMNTCVENKKESKKLLE
53	1403	A	1011	1	630	PEVIQQSAYDSKADIWSLGITAIELAKGEPPNS
] "	1703] ^	1011	•	0.50	DMHPMRVLFLIPKNNPPTHCWRRLLESFKEV
	ł	1		j i		*LMLA*TKDPSI\RPTAKELLKHKFIVKNSKKT
1	Ì	1				SYLTELIDRFKRWKAEGHSDDESDSEGSDSES
J	J	}				TSRENNTHPEWSFTTVRKKPDPKKVQNGAEQ
ļ - · · ·	ļ·· -·					DLVQTLSCLSMITPAFAELKQQDENNASRNQ
1	l	i		İ		AIEELEKSIAVAEAAGPG
54	1404	A	1016	1	222	ISIDA*KAFDKIOH/CFMITTLKKLGIDGKYLN
1		ļ ^^	1210	•	عدد	TIKAIDDRHTVSTILNVEKLKAFL*RSGTRQRF
İ						PISGSGARI
55	1405	A	1033	3	366	HASVDGDEGSDDVYYYYTPAILRELQALNTA
1 ""		١.,			200	EAAEHRPEEDRMLSEDPWRPAIIMIKGYMPL
1	1					HNIPHTEVIDVTGLNQSHLYQHLNKGTPMKT
1	l	l				QKRAA\LYTWHVLEQLEILRQINQQSHGPG
56	1406	<u> </u>	1044	5	429	
30	1400	^	1044	ا ا	429	SVLTLQTRSPSKPLS\RKLMDWEVVSRNSISE DRI ETOSPASPSPRVTPNOSOETPVDGVBLAL
		l				DRLETQSRASRSPPVTPNQSQETPVDGKPLAL
1	}	1				PPNQSQKNIRYHIHYLHLQYYLDRHISATLPIP
1		l				SSSGIPTPIAVITDALTDLVELILGQPCSEESGR
1	L	<u></u>	l			APGTLFLLAL

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GAYAFETNGFPIMLVLTTDKIEGDVGIAGLYD
						MH\ISLPMAFLLRTLVRCTSY\IPVTHVLSTPV TCLRRREKDGV\IVDVL\SDTA\S\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\\
58	1408	A	1058	258	419	LKHRDTPVVGANNRALSCTPLTSLTLCALCPL PCLGCPTXATCRLYQTTVAVVF
59	1409	A	1064	3	425	KAFSFTTSLIGHQRMHTGERPYKCKECGKTF KGSSSLNNHQRIHTGEKPYKCNECGRAFSQC SSLIQHHRIHTGEKPYECTQCGKAFTSISRLSR HHRIHTGEKPFHCNECGKVFSYHSALIIHQRIH TGEKPYACKDVGK
60	1410	A	1065	204	419	GGPPGPFLAHTHAGLQAPGPLLAPAGDEGDL LLLAVQQSCLADHLLTASWGGK/DPIPTKALG EGQEGLPLTV
61	1411	A	1079	3	383	RHSRAHLCQPFHLVMRDLLQLGQDIPQGCHY LEENHLIHRDIAARNCLLSCAAPTRAATIGDF GMARYIYRTRYYQLGDRAL/LPRKWMPPEAL LEGIFTYNTDSWTFGVLLWEIFSLGYMPYPGR TN
62	1412	A	1080	1	859	VVEFLWSRRPSGSSDPRPRRPASKCQMMEER ANLMHMMKLSIKVLLQSALSLGRSLDADHA PLQQFFVVMEHCLKHGLKVKKSFIGQNKSFF GPLELVEKLCPEASDIATSVRNLPELKTAVGR GRAWLYLALMQKKLADYLKVLIDNKHLLSE FYEPEALMMEEEGMVIVGLLVGLNVLDANL\ CLKGEDLDSQVGVIDFSLYLKDVQDLDGGKE HERITDVLDQKNYVEELNRHLSCTVGDLQTK IDGLEKTNSKLQERVSAATDRICSLQEEQQQL REQNELIR
63	1413	A	1083	2	615	SSFAKHKRIHTGEKPFICLECGKAFTSSTTLTK HRRIHTGEKPYTCEECGKAFRQSAILYVHRRI HTGEKPYTCGECGKTFRQSANLYAHKKIHTG EKPYTCGDCGKTFRQSANLYAHKKIHTGEKP YKCKECGKAFKSYYSILKHKRTHTRGMSYEG DEC/QRSLN/RSSILSNHKIIHNEEK/PLKCEKCE KAFNHTSICCRHKKN
64	1414	A	1084	946	1	KKQDLSSSLTDDSKNAQAPLALTESHLATLA SSSQSPEAIKQLLDSGLPSLLVRSLASFCFSHIS SSESIAQSIDISQDKLRRHHVPQQCNKMPITAD LVAPILRFLTEVGNSHIMKDWLGGSEVNPLW TALLFLLCHSGSTSGSWNLGAQQDQCKISFS FFSWLTTGLTTQQRTAIENATVAFFLQCISC HPNNQKLMAQVLCELFQTSPQRGNLPTSGNI SIGFIRIRLFLQLMLEDEKVTMFLQSPCPLYKG RINATSHVIQHPMYGAGHKFRTLHLPVSTTL SDVLDRVSDTPSITAKLISKQKDDKKKK
65	1415	A	1087	103	324	PRAFEFVHTEMIVG/RVQNIHLFTLQVLEDRA LFTMSVGSSLWSTYLIHVMALP/DRELLKPNA SVALHKLSNALV
66	1416	A	1095	3	493	HETCSVTHIVSFSLPFLNPSHPASTPGHTENEQ PSLVWFDRGKFYLTFEGSSRGPSPLTMGAQD TLPVAAAFTETVNAYFKGADPSKCIVKITGE MVLSFPAGITRHFANNPSPAALTFRVINFSRLE HVLPNPQLLCCDNTQNDANTK\EFWVNMPNL MTHLK
67	1417	A	1098	57	356	LKLTSLGFIIGVSVVGNLLISILLVKDKTLHRA PYYFLLDLCCSDILRSAICFPFVFNSVKNGST WTYGTLTCKVIAFLGVLSCFIITAFMLFCISVT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RYL
68	1418	A	1106	•	1326	MGKISATGINMGTKCSWALVWHLESYDPKH YEREGMQDWKTASGQSEEATQQSSQKPQPH YTTYQSSSFLKYSSESHLLAWRENSSEGSFQF PGRSRARPPRTRQQRRGAAAGPGRGAVRLG HPQSAAQPQLRAAARIPESPAAFPAQPRPGSA RNSDASGPASLSRTLGRASSPRPPQAPDVTAP SPAALAPRAARGGSRAAALAGAEAEEPLRTL APRPTRAAAPPPPPPPPPLPPGAPPPPVRCVSR RARAPPWR/PAATGPPPRPVAPSRKLGSARAP APALQIRKGTSSGLPGRGGGSGPGNNLSSVA GNWRGSSFAVERPGMAKYQGEVQSLKLDDD SVIEGVSDQVLVAVVVSFALLATLVYALFRNV HQNIHPENQELVRVLREQLQTEQDAPAATRQ QFYTDMYCPICLHQASFPVETNCGHLFCGSLT PNSIW
69	1419	A	1107	2	466	FDTARLHEFGTSITQIFAVDNREDLQKWMEA FWQHFFDLSQWKHCCEELMKIEIMSPRKPPLF LTKEATSVYHDMSIDSPMKLESLTDIIQKKIEE TNGQFLIGQREESLP/SS/CGPHSLMVTIKWSS RKRY/SYPASEPLHDEKGKKRQAPLPPSDK
70	1420	A	1111	698	23	ALRRLHYVRATKVVFLSFRRPFWREEHIEGGH SNTDRPSRMIFYPPPREGALLLASYTWSDAAA AFAGLSREEALRLALDDVAALHGPVVRQLW DGTGVVKRWAEDQHSQGGFVVQPPALWQT EKDDWTVPYGRIYFAGEHTAYPHGWVETAV KSALRAAIKINSRKGPASDTASPEGHASDMEG QGHVHGVASSPSHDLAKEEGSHPPVQGQLSL QNTTHTRTSH
71 .	1421	A	1119	2	385	QKQTLQNGYLDSSMDÏLYLGSLPPELQVSSDE PPGPPEQAGLSQFHLEPETQNPETTEEIQSS\LQ QEAAAQLPQLPEVVELSSTKA\EAPALPSQSL EGVHSSTEQKAPAQQLPAFEEILAPLLIHHE
72	1422	A	1127	1	906	HAQYVGPYRLEKTLGKGQTGLVKLGVHCIT GQKVAIKIVNREKLSESVLMKVEREIAIL\RLI EHPHVLKLHGVYENKKYFPPDELTSGPSMLA QVSPHGKLSARRSWDLLSGFPRYLVLEHVSG GELFDYLVKKGRLTPKEARKFFRQIVSALDFC HSYSICHRDLKPENLLLDEKNNIRIADFGMAS LQVGDSLLETSCGSPHYACPEVIKGEKYDGR RADMWSCGVILFALLVGALPFDDDNLRQLLE KVKRGVFHMPHFIPPDCQSLLRGMIEVEPEKR LSLEQIQKHPWYLGGNFIS
73	1423	A	1128	1	802	LRNALDVLHREVPRVLVNLVDFLNPTIMRQV FLGNPDKCPVQQA/MLEPLGSKTETLDLRAE MPITCPTQNEPFLRTPRNSNYTYPIKPAIENWG SDFLCTEWKASNSVPTSVHQLRPADIKVVAA LGDSLTTAVGARPNNSSDLPTSWRGLSWSIG GDGNLETHTTI.PNILKKFNPYLLGFSTSTWEG TAGLNVAAEGARARDMPAQAWDLVERMKN SPDINLEKDWKLVTLFIGGNDLCHYCENPEA HLATEYVQHIQQALDILSE
74	1424	A	1139	60	480	FREPCLLVPGDHQPLREASWLA/LPPIGLWGT DSPLCCVEVAIPCNKGAHSVGLKGWLLAQG VLGMRDTIPQEHPWESTPDLCFCRDPEEIEVE EQPAADAAVAKGEF/QGEQIAPVPA\IIAAHPE AADPAPVHTTAHPKGA
75	1425	A	1147	2	413	PFPHQHPQEPKGSCWPQSALRGQCPGPVLGV TTTSDLCSLQVPVSSHRNPLLDLAAYDQEGR

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RFDNFSSLSIQWESTRPVLASIEPELPMQLVSQ DDESGQKKLHGLQAILVHEASGTTAITATAT GYQESHLSSAR PIISAPAQDDPILLSFIHCLHANLLCVWRRDVK PDCKEIWIFWWGDEPNLVVOYIMNCMLWK
<u> </u>	3					KDSGKMAFPMNVGRC/FFKEIHNLLERCLMD KNFVLIGKWFVRPYYKDEKPVNKSEHLSCAF T
77	1427	A	1162	526	350	RFPQGLEDVSTYPVLIEELLSRGWSEEELQGV LRGNLLRVFRQVEKVQEENKWQSPLED
78	1428	A	1171	I	1293	MAESASPPSSSAAAPAAEPGVTTEQPGPRSPP SSPPGLEEPLDGADPHVPHPDLAPIAFFCLRQT TSPRNWCIKMVCNPWFECVSMLVILLNCVTL GMYQPCDDMDCLSDRCKILQVFDDFIFIFA MEMVLKMVALGIFGKKCYLGDTWNRLDFFI VMAGMVEYSLDLQNINLSAIRTVRVLRPLKA INRVPSMRILVNLLLDTLPMLGNVLLLCFFVF FIFGIIGVQLWAGLLRNRCFLEENFTIQGDVAL PPLYYQPEEDDEMPFICSLSGDNGIMGCHEIPP LKEQGRECCLSKDDVYDFGAERQDLNASGL CVNWNRYYNVCRTGSANPHKGAINFDNIGY AWIVIFQVITLEGWVEIMYYVMDAHSFYNFI YFILLIIVSVREPGLLGGSFSTAQSPKCQGDSFP GVAAESLLLRGWVLWLPGGG
79	1429	A	1175	1	405	PNDFFKDMFPDLPGGPLGPIKAENDYGAYLN FLSATHLGGLFPPWPLVEERKLKPKASQQCPI CHKVIMGAGKLPRHMRTHTGEKPYMCTICE VRFTRQDKLKIHMRKHTGERPYLCIHCNAKF VHNYDLKNHMR
80	1430	A	1182	25	198	EMNELSQQLSQQGGRGASQCPSPPAPTLPNPT PLCQLQLQRVNTGLPTPPCHPGAGAA
81	1431	A	1186	254	583	KTVLDVGAGTGILSIFCAQAGARRVYAVEAS AIWQQAREVVRFNGLEDRVHVLPGPVETVEL PEQVDAIVSEWMGYGLLHESMLSSVLHARTK VVKDGGFFLPXSSELFM
82	1432	A	1187	2	716	DFVDAARNLPLESTKSPAEPSKSVPSLEVDPRA SSQGLPSQGPVQNQGRRGEQRPKKF/TVIQHT SSFEKSDSLEQPSGLEGEDKPLAQFPSPPPAPH GRSAHSLQPKLVRQPNIQVPEILVTEEPDRPD TEPEPPPKEPEKTEEFQWPQGSQTLAQFPVEK LPPKKKRLGLAKMAQSSGESSFESSVPLFRSP SQESNVSLSGSSRSALFERDDHGKAEAPSPSF DMGPKPLGTHMLTV
83	1433	A	1188	517	804	ESPGLSKVLRTGAFAYPFLFDNLPLFYRLGLC WGRGHGCGQEALSTSHGYHLFCALLTGFLFA- SHLPERLAPGRFDYIGHSHQLFHICAVLGTHF Q
84	1434	A	1192	45	476	LGDVGFWVERTPVHEAAQRGESLQLQQLIES GACVNQVTVDSITPLHAASLQGQARCVQLLL AAGAQVDARNIDGSTPLCECLRLGQHRVCEA LAVLRGQGQPSPVHSVPPARGLHXREFRMC* GFLFDVGXNLEAHEFHFGEP
85	1435	A	1194	69	410	KRSEEASAPPFPLGGTGAAPTRASLPEQILLPR SCLEARKSQPDEKLLSALHNSRTWN*EPRRSQ HRLVSPEVHPGRRGSSPGVAECKLTSAYFRT GRSPCPSLPGTTRTNSLL
86	1436	A	1215	3	405	LPSHTCGNPGRLPNGIQQGSTFNLGDKVRYSC NLGFFLEGHAVLTCHAGSENSATWDFPLPSC RADDACGGTLRG/AEWHHLQPPLPLG/ATKN

Ceromo	CEATE	114-4	Tero	I n 3*	1 W 10 10 10 10 10 10 10 10 10 10 10 10 10	
SEQ ID NO: of	SEQ ID NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in in	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
"	1	ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ	j		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	j	}	1	peptide		/-possible nucleotide deletion, \-possible
			1	sequence		nucleotide insertion
						NADCTWTILAELGDTIALVFIDFQLEDGYDFL
			i			EVTGTEGSSLW
87	1437	A	1216	226	964	GTARFGPMVGFGANRRAGRLPSLVLGVLLV
		İ				VIVVLAFNYWSISSRHVLLQEEVAELQGQVQ
}		Į				RTEVARGRLEKRNSDLFAVVGHAQETDRPEG
		Ì	1			GRLRPPQQPAAGQRGPREEM\EDDKVKLQNN
]	J			ISYQMADIHHLKEQLAELRQEFLRQEDQLQD
						YRKNNTYLVKRLEYESFQCGQQMKELRAQH
						EENIKKLADQFLEEQKQETQKIQSNDGKELDI
		ļ				NNQVVPKNIPKVAENVADKNEEPSSNHIPHG
88	1438	A	1218	1	534	PEFGTTISCGYLMATDVSRRPSVHKAVEIEQE
						RVKSAGAWIIHPYSDFRFYWDLIMLLLMVGN
						LIVLPVGITFFKEENSP\PWIVFNVLSDTFFLLD
			[LVLNFRTGIVVEEGAEILLAPRAIRTRYLRTW
						FLVDLISSIPVDYIFLVVELEPRLDAEVYKTAR
89	1439	A	1223	1	743	ALRIVRFTKILSLLRL
0.7	1439	A	1223	1	/43	MGFDEVFMINLRRRQDRRERMLRALQAQEIE
						CRLVEAVDGKVGMLTRSNAAPGRHLAMLET
				, ,	1	LVVVAPRFVDADNLILNPDTLSLLIAENKTVV
						APMLDSRAAYSNFWCGMTSQGYYKRTPAYI PIRKRDRRGCFAVPMVHSTFLIDLRKAASRNL
	i ,					VAFYPPHPDYTWSFDDIIVFAFSCKQ\AEVQMY
						VCNKEEYGFLPVPLRAHSTLQDEAESFMHVQ
				1		LEVMVPSSPSSAQSMAVVSADHIGLVISYL
90	1440	Α	1227	2	349	NKTSFIFYLKNIVVADLIMTLTFPFRIVHDAGF
				i		GPWDFKFILCRYTSVLFYANMDTSIVVLGLIT/
					ļ	YDRY/WKVVRHL/WDSWMTGI/SFTRVYLLG
						LGARLVWFGKLILAKGGHGGISWL
91	1441	A	1245	3	1937	LGSSDVRAPQRSELGAESPSRMVASQAYNLT
	l			l		SALTPILTRSRVLNEEPLTLAGF\SRAPANLSD
		i		ł	}	VVQLIFLVDSNPFPFGYISNYTVSTKVASMAF
l				l		QTQAGAQIPIERLASERAITVKVPNNSDWAAR
}						GHRSSANSVVQPQAFVGAVVTLDSSNPAAV
١ .		j	J	İ	[LHLQLNYTLLDGRYLSEEPEPYLAVYLHSEPR
ļ						PNEHNCSASRRIRPESLQGADHRPYTFFISPGT
		į	1	j		RDPVGSYRLNLSSHFRWSALEVSVGLYTSLC
ĺ	l	-	ľ			QYFSEEDVVWRTEGLLPLEETSPRQAVCLTR
						HLTAFGTSLFVPPSHIRFVFPEPTADVNYIVML
1	}					TCAVCLVTYMVMAAILHKLDQLDASRGRAIP FCGOPGPFKYEII VKTGWGPGSGTTAHVGDA
ł		i	1		}	FCGQRGRFKYEILVKTGWGRGSGTTAHVGIM LYGVDSRSGHRHLDGDRAFHRNSLDIFOIATP
	İ	- 1	i			HSLGSMWKIRVWHDNKGLSPAWFLOHIIVRD
		J	1		J	LQTARSTFFLVNDWLSVETEANGGLVEKEVL
,		ļ			1	AASKASFRVPTPS\AALLRFRRLLVAELORGF
		1				FDKHIWLSIWDRPPRSCFTRIQRATCCVLLICL
i	1	ľ	l		1	FLGANAVWYGAVGDSAYSTGRVSRLNPLSV
- 1	1	1	i	1	1	DTVAVGLVSSVVVYPVYLAILFLFRMSRSKV
1		ļ			Ì	GWGWGPGSTGNGAWASAPCPEPPLSSAAAR
		1				GKGVHQRLLGKGQHT
92	1442	A	1246	5	562	VFDEENILNELNDPLREEIVNFNCRKLVATMP
ì		l]			LFANADPNFVTAMLSKLRFEVFQPGDYIIREG
	j	j	ŀ	ĺ	ĺ	AVGKKMYFIQHGVAGVITKSSKEMKLTDGS
				1	ĺ	YFGEICLLTKGRRTASVRADTYCRLYSLSVD
	1	l	j		l	NFNEVLEEYPMMRRAFETVAIDRLDRIGKKN
اـــــا						SILLQKFQKDLNTGVFNNQENEILKQIVKH
93	1443	A	1249	180	901	TVPPPPGGPSPAPLHPKRSPTSTGEAELKEERL
- 1	1	}	ļ	ļ	}	PGRKASCSTAGSGSRGLPP\SSPMVSSAHNPN
1	l	ł	- 1	1	i	KAEIPERRKDSTSTPNNLPPSMMTRRNTYVCT
	i	- 1	i			ERPGAERPSLLPNGKENSSGTPRVPPASPSSHS

NO: of No: of nucleotide octicle sequence Page	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino agid sogrence (A=A1
nucleotide cotide cotide corresponding of personal properties of per						1	Amino acid sequence (A=Alanine C=Cysteine,
Contragnoding Sequence USSN Solution Sequence International International Int			1.00				E-Phenylalanine G-Glucina U-Uistidia
sequence Sequence Solidar Sequence		1	1	1			
But But		1 -	Ì	1			
amino acid residue of peptide sequence p		dence	}	1			
Peptide Sequence	ucita	ì	ł	314 /			
Popside Popside Popside Popsible nucleotide insertion Popsible nucleotide insertion Popsible nucleotide insertion Popsible nucleotide LAPPSGERSILARGSTIRSTFHGGQVEDRRAK CGGGGGVONDPASPTLAHEAAPIPAGRPRI TINLFTKLTSKLTRRV ADPERIGGEVTRR ROEDHLSPGGRCSEL ROEDHLSPGGRCSEL ROEDHLSPGGRCSEL ROEDHLSPGGRCSEL ROEDHLSPGGRCSEL Popsible	l	1					1=1 hreonine, V=Valine, W=Tryptophan,
	1	1	ł	ļ		sequence	Y=1 yrosine, X=Unknown, *=Stop codon,
LAPPSGERSILARGSTIRSTFHGGQVEDRRAK GGGGGVQNGPPASPTLAHEAAPLPAGRPRI TINLFTKLTSKLTRRVADEPERIGGEVTRRP RQEHLISPGGRGCSEL ROBWILTPKLSKLTRRVADEPERIGGEVTRRP RQEHLISPGGRGCSEL ROBWILTPKLSNASPWISLVKKLMKKW VTQNLTTREQLEAGRYFDLRVSSKPGDADQ EIVFHIGLFIGKWDGLMEDESICLOPH/DEUEI DFNHFYAMDETHHKCLVLRIQEAFGNKLCP/CR	1	ļ				}	
94 1444 A 1261 3 385 KFSOWGLTEKELKENSAPMISLVEKKLMKKW 95 1445 A 1282 2 550 GPRDNFGIEDRIFIGAMFIFELVARF VAPPILIFERALINIAL ANTIFELLIA AND THE STEEL STATEMENT OF THE STATEMENT OF T	<u> </u>	 _	↓		sequence		
P4 1444 A 1261 3 385 KFSQWGLTKPKLSKLTKKVADEPERIGGEVTREP RQEDHLSPGGRGCSEL SEGWGLTKPKLSNASPWJSLVKKLMKKWS VTQNLTFREQLEAGIRYFDLRVSKRGADQ ETYFHIGLFGIKVWDGLMEDSELTQH-QELIF DPNHFYAMDETHHKCLVLRIQEAFGNKLCP/CR R P5 1445 A 1282 2 550 GPRDMPGEDPRFEIVEHFGIAWFTFELVARF/ VAPPELKFRNALNLDLMSIVPFYTILVVAL VVESTPTLANLGRVAQVLRLMRIFRILKLARF STGLRSLGATLKYSYKEVGLLLLVLSVGISTE VVAYTIEKEENEGLATIRACWWATVSMTI VQYGDVVPGTTAGKLTASACILA PVAYTIEKEENEGLATIRACWWATVSMTI VQYGDVVPGTTAGKLTASACILA STGLRSLGATLKYSYKEVGLLLLVLSVGISTE VVAYTIEKEENEGLATIRACWWATVSMTI VQYGDVVPGTTAGKLTASACILA STGLRSLGATLKYSYKEVGLLLLVLSVGISTE VVAYTIEKEENEGLATIRACWWATVSMTI VQYGDVVPGTTAGKLTASACILA STGLRSLGATLKYSYKEVGLLLVLUSUGISTE VVAYTIEKEENEGLATIRACWWATVSMTI VQYGDVVPGTTAGKLTASACILA STGLRSLGATLKYSYKEVGLLLVLUSUGISTE VVAYTIEKEENEGLATIRACWWATVSMTI VQYGDVVPGTTAGKLTASACILA AETTYMIFLVILLUSLYLLULLUSLYLLULLAVVAMA YEQNQATLEEAGOKEAFGQMEGUKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEA KLSSKSAKERNRRKKRKQKEQSGGEEKDE EFQKSSESDSTRRGGFRSFEGNLTVEKKYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTEDNESRRDSLFYPRHGI RNNSNLSJTSRSSRMLAVFFTANGKMHSTVDC NGVVSLVGGPSVFTSPVGQLLFVEUIDKPATD DNGTTTETEMFKKRSSSFPNANGTLTVEKKYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTEDNESRRDSLFYPRHGI RNNSNLSJTSRSSRMLAVFFTANGKMHSTVDC NGVVSLVGGPSVFTSPVGQLLFVEUIDKPATD DNGTTTETEMFKKRSSSFPNANGTHSTPTGPCGALS SELPSTSPSVAAISSRSPNANFTSIEDPOGR SELPSTSPSSVAAISSRSVHKFFTQSRIPPLD MIPPAPRHITTEELSVLESCLIRKWTEIENDTR DLQESISRIHKTIELMYSDKSMIQVPYRLHAV LVHGQQANAGHWANTFDHFESR WMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAGSTRUFFTDYRLHVVVYYGNQ APKKIEKTLEGFGDRNLSPERCHNIMKVA QKALEMIKPEEVIDSLEFVEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLLCAYQNIKELL SKGLYRGHDEELISHYRRECLILKLNEQAAELIF ESGEDREVNINGLIMNIFFIVFFILLLUKDEQAAELIF ESGEDREVNINGLIMNIFFIVFFILLUKDEMEE SGEDREVNINGLIMNIFFIVFFILLUKDEMEE SGEDREVNINGLIMNIFFIVFFILLUKDEMEE SGEDREVNINGLIMNIFFIVFFILLUKDEMEE SGEDREVNINGLIMNIFFIVFFILLUKDEMEE	1	ĺ	į.	1	1	ŀ	LAPPSGERSRLARGSTIRSTFHGGQVRDRRAG
P4 1444 A 1261 3 385 KFSQWGLTKPKLSKLTKKVADEPERIGGEVTREP RQEDHLSPGGRGCSEL SEGWGLTKPKLSNASPWJSLVKKLMKKWS VTQNLTFREQLEAGIRYFDLRVSKRGADQ ETYFHIGLFGIKVWDGLMEDSELTQH-QELIF DPNHFYAMDETHHKCLVLRIQEAFGNKLCP/CR R P5 1445 A 1282 2 550 GPRDMPGEDPRFEIVEHFGIAWFTFELVARF/ VAPPELKFRNALNLDLMSIVPFYTILVVAL VVESTPTLANLGRVAQVLRLMRIFRILKLARF STGLRSLGATLKYSYKEVGLLLLVLSVGISTE VVAYTIEKEENEGLATIRACWWATVSMTI VQYGDVVPGTTAGKLTASACILA PVAYTIEKEENEGLATIRACWWATVSMTI VQYGDVVPGTTAGKLTASACILA STGLRSLGATLKYSYKEVGLLLLVLSVGISTE VVAYTIEKEENEGLATIRACWWATVSMTI VQYGDVVPGTTAGKLTASACILA STGLRSLGATLKYSYKEVGLLLLVLSVGISTE VVAYTIEKEENEGLATIRACWWATVSMTI VQYGDVVPGTTAGKLTASACILA STGLRSLGATLKYSYKEVGLLLVLUSUGISTE VVAYTIEKEENEGLATIRACWWATVSMTI VQYGDVVPGTTAGKLTASACILA STGLRSLGATLKYSYKEVGLLLVLUSUGISTE VVAYTIEKEENEGLATIRACWWATVSMTI VQYGDVVPGTTAGKLTASACILA AETTYMIFLVILLUSLYLLULLUSLYLLULLAVVAMA YEQNQATLEEAGOKEAFGQMEGUKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEA KLSSKSAKERNRRKKRKQKEQSGGEEKDE EFQKSSESDSTRRGGFRSFEGNLTVEKKYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTEDNESRRDSLFYPRHGI RNNSNLSJTSRSSRMLAVFFTANGKMHSTVDC NGVVSLVGGPSVFTSPVGQLLFVEUIDKPATD DNGTTTETEMFKKRSSSFPNANGTLTVEKKYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTEDNESRRDSLFYPRHGI RNNSNLSJTSRSSRMLAVFFTANGKMHSTVDC NGVVSLVGGPSVFTSPVGQLLFVEUIDKPATD DNGTTTETEMFKKRSSSFPNANGTHSTPTGPCGALS SELPSTSPSVAAISSRSPNANFTSIEDPOGR SELPSTSPSSVAAISSRSVHKFFTQSRIPPLD MIPPAPRHITTEELSVLESCLIRKWTEIENDTR DLQESISRIHKTIELMYSDKSMIQVPYRLHAV LVHGQQANAGHWANTFDHFESR WMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAGSTRUFFTDYRLHVVVYYGNQ APKKIEKTLEGFGDRNLSPERCHNIMKVA QKALEMIKPEEVIDSLEFVEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLLCAYQNIKELL SKGLYRGHDEELISHYRRECLILKLNEQAAELIF ESGEDREVNINGLIMNIFFIVFFILLLUKDEQAAELIF ESGEDREVNINGLIMNIFFIVFFILLUKDEMEE SGEDREVNINGLIMNIFFIVFFILLUKDEMEE SGEDREVNINGLIMNIFFIVFFILLUKDEMEE SGEDREVNINGLIMNIFFIVFFILLUKDEMEE SGEDREVNINGLIMNIFFIVFFILLUKDEMEE	1	i	1	ł		1	GGGGGGVQNGPPASPTLAHEAAPLPAGRPRP
P4	ì	Į.	l	1	ĺ	Ì	TINLFTKLTSKLTRRVADEPERIGGPEVTRRP
94 1444 A 1261 3 385 KFSQWGLTKPKLSNASPWISLVKKLMKKW VTONLTREGLEAGRYFDLYSKSKPOADO EIYFHGLFGIKVWDGLMEIDSFLTQHPQEIFI DPNHFYAMDETHHKCLVLRIQEAFGNKLCP/ CR 95 1445 A 1282 2 550 GFRDNPGEDPREIVEHFGIAWFTFELVARF/ VAPPELKFKNALNLDLMSIVPPYTLVVNLL VVESTPTLANLGRVAQVLRLMRIPELKLAR STGLRISGATLK YSYKEVGLLLLYLSVGISFR VVAYTIEKEENEGLATIFACWWWATVSMTI VQYGDVPGTTAGKLTASACILA 96 1446 A 1294 1 1456 QLLPPSNREMAGLIVGRCLGSAALRPVGDLIT SSGQVAVNAPQAGSAAGKGKFQDNFERG YYKKFFDANCHEDVDYNPAGQGGFTEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMIFILVLVILLGSLYLVTLLLAVVAMA YEEQNQATLEEAFGKEAFFQQMLEGQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEA KLSSKSAKERRNRKKKKQKEQSGGEEDE EFQKSSEEDSIRRGGFPSFEGNRLTYERKYSS KLSSKSAKERRNRKKKKKQKEQSGGEEDE EFQKSSEEDSIRRGGFPSFEGNRLTYERKYSS KLSSKSAKERRNRKKKKKKQKEQSGGEEDE EFQKSSESSIRRGGFPSFEGNRLTYERKYSS RNSNLSQTSRRSGFPSFTSVGQLLFVEUNDKPATD DNGTTTETEMFKKRSSYPVANGKHBSTVDC NGVVSLVGGPSVFTSPVGQLLFVEUNDKPATD DNGTTTETEMFKKRSSYPVANGKHBSTVDC NGVVSLVGGPSVFTSPVGQLLFVEUNDKPATD DNGTTTETEMFKKRSSSYPTANGKMHSTVDC VGTSPVDIDDASSPPSGBPSGULFVEUNDKPATD DNGTTTETEMFKKRSSSYPTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPDL MIPPARHITTEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYLHAV LVHEGQANAGHWAYTPIELBERDTR DLQESISRIHRTIELMYSDKSMIQVPYLHAV LVHEGQANAGHWAYTPIELDRFV EEDNQRFEKELEEWDAQLAKAGLGKLLAS QKLRESSTSVTTAQAAGDPKYLEOPSRSPFSK HLKEETIQITIKASHEHEDISPOERCHNIMKVA QKLRESSTSVTTAQAAGDPKYLEOPSRSPFSK HLKEETIQITIKASHEHEDISPTUQSAIKLE YARLVKLAQEDTPPETDVRLHHVVYVYIQNQ APKKIEKTLLEQFGDRNLSPERCHNIMKVA QAKLEMIKPEGVBLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLLCAYQNIKELL SKGLYRCHDELISHYRRECLLKLKDEQAAELF SGEBREVNNGLIMNFFIVPFLILLLVDEMEE	L	1	L		<u> </u>		RQEDHLSPGGRGCSEL
PS 1445 A 1282 2 550 GPRDNPG\(\text{CR}\) 95 1445 A 1282 2 550 GPRDNPG\(\text{CR}\) 96 1446 A 1294 1 1456 QLLPPS\(\text{CR}\) 97 1446 A 1294 1 1456 QLLPPS\(\text{CR}\) 98 1446 A 1294 1 1456 QLLPPS\(\text{CR}\) 98 1446 A 1294 1 1456 QLLPPS\(\text{CR}\) 99 1446 A 1294 1 1456 QLLPPS\(\text{CR}\) 90 1446 A 1294 1 1456 QLLPPS\(\text{CR}\) 91 1456 QLLPPS\(\text{CR}\) 92 1446 A 1294 1 1456 QLLPPS\(\text{CR}\) 93 1447 A 1295 2 2057 QFR\(\text{CR}\) 94 1447 A 1295 2 2057 QFR\(\text{CR}\) 95 1447 A 1295 2 2057 QFR\(\text{CR}\) 96 1447 A 1295 2 2057 QFR\(\text{CR}\) 97 1447 A 1295 2 2057 QFR\(\text{CR}\) 98 1447 A 1295 2 2057 QFR\(\text{CR}\) 99 1447 A 1295 2 2057 QFR\(\text{CR}\) 90 1447 A 1295 2 2057 QFR\(\te	94	1444	A	1261	3	385	KFSQWGLTKPKLSNASP/WISLVKKLMKKWS
BIYFHIGLFGIKVWDGLMEIDSFLTOHPOEUE DPNHFYAMDETHHKCLVLRIQEAFGNKLCPA CR]	j	ļ	,		1	VTONLTFREOLEAGIRYFDLRVSSKPGDADO
95 1445 A 1282 2 550 GPRDNPGVEDPRFEIVEHFGIAWFTFELVARFA VAPDFLKFFKNALNLIDLMSVPPTTILVVNL VVSSTPTLANLGRVAQVIRLMRIFIRLKLAR STGLRSLGATLKYSVKEVGLLLLVLSVGISFS VVAYTIEKEENEGLATIPACWWWATVSMTI VGYGDVVPGTTAGKLTASACILA 96 1446 A 1294 1 1456 QLLPPSNRENAGILVGRCLCSAALRPVGDLTI SSGGVAVRNAPQAGSAKGKKPQDDNFEFIC YFKKFFDANCNEKDYNPVAAGQGQETEVAP SIVAPVLNKPNQCPEGYTCVKAGRNNYGYT SPDTTSWAFLSLFRLMTDOYWENLYQLTLRA AETTYMIPLVIVULLGSLYLVTLILAVVANM YEEQNQATLEEAGVEAFCQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEA KLSSKSAKERNRIKKKRQKEQSGGEKDE FFQKSESEDSIRRKGFREGONLFYCKRYS PHQSLLSIRGSLFSPRRNSRTSLFSFGGRAKDV GSENDFADDEHSTFEDNESSFNSKFTVPRCHGE RNNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIDKPATD DNGTTTETEMBKRRSSSFVSMDFLEDPSQR RRNSSLSGTSRSSRNAVPFANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIDKPATD NGYVSLVGGPSVPTSPVGQLLPEVIDKPATD NGYVSLVGGPSVPTSPVGQLLPEVIDKPATD NGYVSLVGGPSVPTSPVGQLLPEVIDKPATD DNGTTTETEMBKRRSSSFVSMDFLEDPSQR CRAMSLASLITNIVE 97 1447 A 1295 2 2057 IQTQLPTKSQQLKGGNCVRCKMQMNFIAE EVLLKYRTIFYNNNKGPNMLYIEIKAFVHFM NRYLSYGSGPKRFFLVDVLQVALEFASSKPV CTSPVDDIDDASSPPSGSDTJPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHTEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSGNQVPYRLHAV LVHEGQANAGHYWAYIFDERESRWMXYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQELIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALGEKLLAS OKLRISETSVTTAQAAGADFKYLGOFSRSDFSK HLKEETIQITIKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDIPPETDYRLHHVVVYFIQNQ APKKIEKTLLEQFGDRNJEECHNINKVA QAKLEMIKPEEVNLEEVEEWHQDYKFREET MYLIGLENPORSSYNDSLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF BSGEDREVNNGLIIMMEFIPPFLILVDEMEE BSGEDREVNNGLIIMMEFIPPFLILVLDEMEE BSGEDREVNNGLIIMMEFIPPFLILLDEMEE BSGEDREVNNGLIIMMEFIPPFLILLDLEDEFCHNINKVA QAKLEMIKPEEVNLEEVEEWHQDYKEPLE BSGEDREVNNGLIIMMEFIPPFLILLDLEDEFCHNINKVA QAKLEMIKPEEVNLEEVEEWHQDYKEPL BSGEDREVNNGLIIMMEFIPPFLILLDLEDEFCHNINKVA QAKLEMIKPEEVNIGEREFENDERENDTH MYLIGLENFORSSYNDSLFLICAYQNNKELL BSGEDREVNNGLIIMMEFIPPFLILLDLEDEFCHNINKVA BSGEDREVNNGLIIMMEFIPPFLILLDLEDEFCHNINKVA BSGEDREVNNGLIIMMEFIPPFLILLDLEDEFCHNINKVA BYTHTON TOWALTH TOWALTH TOWALTH TOWALTH TOWALTH TOWALTH TOWALTH TO	1	ļ			l	1	EIYFIHGLEGIKVWDGLMEIDSFLTOHPOEUEI
95 1445 A 1282 2 550 GPRDNPGGEDPRFEIVEHFGIAWFIFELVARFA VAPDFLKFFKNALNLIDLMSIVPFYITLVVNL VVESIPTLANLGRVAQVIRLMRIRILKLARR STGLRSLGATIKYSYKEVGLLLI YLSVGISIR VVAYTIEKEENBGLATIPAGWWATVSMTT VGYGDVVPGTTAGKLTASACILA 96 1446 A 1294 1 1456 QLLPPSNRENAGLLVGRCCSAALRPVGDLTI SSGQVAVRNAPQAGSAKAGKGKFQNDFFEIV FYKKFFDANCNEKDYNPVAAGQQGETEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SPDTFSWAFLSLFRLMTYDDYWENLYQLTLRA AETTYMIFLVLVLULIGS.YLVTLLLAVVAMA YEEQNQATLEEAEQKEAFFQMLDELYKXQO EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRKKRVQKEOSGGEKDEI EPQKSESEDSIRRKGFRSTEGORLTYERRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRASDLYPRRHGI RNSNLSOTSSSSMLAVSTSPSRAKKRVQKEOSGGEKDEI EPQKSESEDSIRRKGFRSTEGORLTYERRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRASDLYPRRHGI RNSNLSOTSSSSMLAVSTSPSNGKMSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMKKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNIVE 97 1447 A 1295 2 2057 QTQLPTKSSQQLRKGGNCVRCKMQMNFILAE EVILLEYRITFYNNNKGRPML YEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPSSSVALSSRSVHKPFTGSRPPDLP MIFPARHTTEEELSVLESCLHRWTEIENDTR DLQESISRIKTIELMYSDKSMQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVEDSFGGYRNASACLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALGEKLLAS QKLRSISETSTSTVTAQAAGOPFYLEOPSRSPSK HLKEETIQIITKASHEHEDKSPETVLQSAKILE YARLVKLAQEDIPPETDYRLHHVVVYFIQNQ APKKHEEKTILLEGFGGRNISTEDRCHNINKVA QAKLEMIKPEEVNLEEVEEWHQDYKRFEETT MYLIGLENPQRSSYDASTERENCHNIKVA QAKLEMIKPEEVNLEEVEEWHQDYKRFEETT MYLIGLENPQRSSYDASTERENCHNIKVA QAKLEMIKPEEVNLEEVEEWHQDYKRFEETT MYLIGLENPQRSSYDASTERENCHNIKVA QAKLEMIKPEEVNLEEVEEWHQDYKRFEETT MYLIGLENPQRSSYDASTERENCHNIKVA QAKLEMIKPEEVNLEEVEEWHQDYKRFEETT MYLIGLENPQRSSYDASTERENCHNIKVA QAKLEMIKPEEVNLEEVEEWHQDYKRFEETT MYLIGLENPQRSSYDASTERENCHNIKVA QAKLEMIKPEEVNLEEVEEWHQDYKRFEETT MYLIGLENPQRSSYDASTERENCHNIKVA QAKLEMIKPEEVNLEEVEEWHQDYKRFEETT MYLIGLENPQRSSYDASTERENCHNIKVA QAKLEMIKPEEVNLEEVEEWHQDYKRFEETT MYLIGLENPQRSSYDASTERENCHNIKVA QAKLEMIKPEEVNLEEVEEWHQDYKRFEETT MYLIGLENPQRSSYDASTERENCHNIKVA QAKLEMIKPEEVNLEEVEEWHQDYKRFEETT MYLIGLENPQRSYDASTERENCHNIKEVA	i		1		1	i	DENHEYAMDETHHKCI VI RIOEAFGNKI CPA
95 1445 A 1282 2 550 GPRDNPGIEDPRFEIVEHFGIAWFTFELVARF/ VAPDFLKFFKNALNLIDLMSIVPYTILVANI VVESTPTLANLGRVAQVIRIMFIRFILKI.AR STGLRSI.GATLKYSYKEVGILLLYLSVGISIPS VVAYTIPKEENBGLATIPACWWATVSMTT VGYGDVVPGTTAGKLTASACILA 96 1446 A 1294 1 1456 QLLPPSNRENAGLLVGRCLCSAALRPVGDLT SGQVAVRNAPQAGSAAGKGKFQDNFEFIG YFKKFPDANCNEKDYNPVAAGQGQETEVAP SIVAPVINFPQCPEGYICVKAGRNPNYGYT SPDTFSWAFLSLFRIMTQDYWENLYQLTILRA AETTYMFFLVI.VILLGSLYLVTILLAVVANA YEEQNQATLEAEAGKAFGQMLEGLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEA KLSSKSAKERRNRKKKQKEQSGGEEKDEL EFQKSESEDSIRRKGEAEFQOMLEGLKKQQ GSENDFADDEHSIFEDNESSREDSIFSFRGARADV GSENDFADDEHSIFEDNESSREDSIFSFRGARADV GSENDFADDEHSIFEDNESSREDSFRYSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTIESSQQLKKGGNCVRCKMQMNFIAE EVLLKYRITFYNNKGPNMLYIEIKAFVHFM NRYLSYGSGFKRFPLVDVLQYALEFASSKPV CTSPVDDDIASSPPSGSPSGTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLRWRTEIENDTR DLQESISRIHRTIELMYSDKSMQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGETLPPDLRDFV EEDNQRFEKELEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIOIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFROM APKKIEKTLLEQFGDRNLSFDERCRNIMKVA QAKLEMIEKTLLEQFGDRNLSFDERCRNIMKVA QAKLEMIKPPEVNLEEYEEWHODYRKFRETT MYLIGGLENFORESVIDSLIFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDDEMEE	1	l	l			!	CR CR
VAPDFLKFFKNALNLIDLMSIVPFYITLVVNL VVESTPILANLGRVAQVLRIMFIRLIKIAR STGLRSIGATLKYSYREVGILLLY LSVGISIR VVAYTIEKEENEGLATIPACWWATVSMTT VGYGDVVPGTTAGKLTASACILA 96 1446 A 1294 1 1456 QLLPPSNRENAGILVGRCLCSAALRPVGDLTI SSGQVAVRNAPQAGSAKAGKGKPQDMFERIC YFKKFFDANCNEKDYNPVAAGQQQETEVAP SIVAPVLNRENQCPBCYICVKAGRNPNYGYT SFDTFSWAFISLFRLMTQDYWENLYQULTLA AETTYMFISLFRLMTQDYWENLYQULTLA AETTYMFISLFRLMTQDYWENLYQULTLA AETTYMFISLFRLMTQDYWENLYQULTLA AETTYMFISLFRLMTGHANAAGRISDSAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA	95	1445	A	1282	2	550	
VVESTPILANLGRVAQVLRIMFIRLIKLAR STGLRSLGATIKYSYKEVGLLLLYLSVGISIFS VVAYTIEKEENEGLATIPACWWWATVSMTI VGYGDVVPGTTAGKLTASACILA 96 1446 A 1294 1 1456 QLEPSINENAGILYGEGLESAALRPVGDLTI SIGQVAVRNAPQAGSAKAGKGKFQDNFEFIC YFKKFFDANCNEKDYNPVAAGGQGETEVAP SIVAPVLNKPNQCPEGYICVKAGRINPYGYT SFDTFSWAFLSLFRLMTQDVWENLYQLTLRA AETTYMIFILVILULUSLYLVTLLLAVVAAMA YEEQNQAATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRISDSSEGAA KLSSKSAKERNRIKKKRKQKEQSGGEEKDEI EPQKSESEDSIRRKGFRSIEGNRLTYEKRYSS PHOSLLSIRGSLFSPRRINSRISTSFRGRAKDV GSENDFADDEHSITEDNESRRDSLFYPRRHOI RRINSNLSQTSRSSRMLAVPPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTIETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTINTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMIFILAE EVLLKYRITFYNNIKGPNMLYIEIKAFVHFMI NRYLSYGSGFKRPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIEINDTR DLQESISRHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIPDHESRWMKYNDID AVTKSSWELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSYTTAQAAGDPKYLEQPSRSDFSK HLKEETIQITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFQNQ AFKKIEKTLLEQFGDRNLSFDERCKNIMKVA QAKLEMIKPFEVNLEEYEEWHODYRKFRETI MYLIGLBENFORESYDISLLFLICAYQNNIKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLLLLVDEMEE ESGEDREVNNGLIIMNEFIVPFLLLLVDEMEE					_	1 -50	
STGLRSLGATLKYSYKEVGLLLLYLSVGISIES VVAYTEKEENEGLATIPACWWATVSMTT VGYGDVVPGTTAGKLTASACILA 96 1446 A 1294 1 1456 QLLPPSNRENAGILVGRCLCSAALRPVGDLTI SSGQVAVRNAPQAGASKAGKGKFQDNFEFIC YFKKFFDANCNEKDYNPVAAGQGCETEVAP SIVAPVLNKPNQCPEGYYICVKAGRNPNYGYT SFDTFSWAFISLFRLMTQDYWENLYQUTLKA AETTYMIF/LV/LVILLGSLYLVTLILAV/VAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQATATATSERSEPSAAGRSLSDSSEAS KLSSKSAKERRNRKKRKQKEQSGEEKDEI EFPGKSESEDSIRRKGFRFSIEGNRLTYEKRYSS PHQSLLSIROSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRHGI RNNSLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLFEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNKGPPMMLYIEIKAFVHPM NRYLSYGGSVERFPLVDVLQYALEFASSKPV CTSFVDDIDASSPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKFFTQSRPPDLP MIPPAPRHTTEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNORFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQFSRSDFSK HLKEETIQIITKASHEHEDKSFETVLQSAIKLE YARLVKLAQEDTPETDVRLHHVVYTFIONQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEVEEWHQDYRKFRETT MYLIIGGLENFORESVEEVHQDYRKFRETT MYLIIGGLENFORESVEEVHQDYRKFRETT MYLIIGGLENFORESVEEVHOLDYNKFRETT MYLIIGGLENFORESVEEVHOLDYNKFRETT MYLIIGGLENFORESVEEVHOLDYNKFRETT MYLIIGGLENFORESVEEVHOLDYNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNINGLIIMMEFITVPFLPLLLUPDEMEE	1	i	(1	
96 1446 A 1294 1 1456 QLLPSSNRENAGILVGRCLCSAALRPVGDLTI SGQVAVRNAPQAGSAKAGKGKFQDNFEFIG YFKKFFDANCNEKDYNPVAAGQGQETEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA ABITYMIFILVI/VILLGSLYLVTLILAV/VAMA YEEQNQATLEAGAKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRKKRKQKEQSGGEKDEL EFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRNSTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVCQLLPEVIIDKPATD DNGTTTETEMEKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGIPSQLTPSTITEQQGALS SELPSTSPSVAAISSRSVHKPFTGSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYAVITDHRESRYMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQITKASHEHEDKSFTVLQSAIKLE YARLVKLAQEDDFSTVDSLIFLICAYONNKELL SKGLYGGHOEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMMEFIVPFLPLLLVDEMEE KDILAVEDMRNIKCSYLGGEMEPPHLOEKLT	•	į.	1	[STCL DOLCATI WYSYKEYOLLUKINKIFKILKLARH
96 1446 A 1294 1 1456 QLLPPSNRENAGLLVGRCLCSAALRPVGDLTI SSGQVAVRNAPQAGSAKAGKGKFQDNFERIÇ YFKKFFDANCNEKDYNPVAAGGGGETEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLKA AETTYMIFILVILVILLGSLYLVTLILAVVAMA YEEQNQATLEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAK KLSSKSAKERNRRKKRKQKEQSGGEKDEL EFQKSESEDISIRKFRFSIEGORILTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRDSLFVPRRHGI RNNSLSGTRSSRALAVPPANGKMHSTVDC NGVVSLVGGFSVPTSPVGQLLPEVIIDKPATD DNGTTTEMEKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTOLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLK YRITFYNNNKGPNMLYIEIKAFVHMI NRYLSYGGSFKRFLVDVLQYALEFASSKPV CTSFVDDIDASSPSGSPSGTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHTEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHVAIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGETLPPDLRBFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSFTVLQSAIKLE YARLVKLAQEDTPPETDVRLHHTVVYFIONQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLNIGLENFQRESYUDSLLFLICAYQNNKELL SKGLYGGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNIGLIIMMEFIVPFLPLLLLVDEMEE	1	i		[
96 1446 A 1294 1 1456 QLLPPSNRENAGLLVGRCLCSAALRPVGDLTI SIGQVAVRNARPAGASSAKAGKGKFQDNFEFIC YFKKFFDANCNEKDYNPVAAGQGOETEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTDDYWENLYQLTLRA AETTYMIFA/V.VILLGSLYLVTLILAVVAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRKKKRKQKGSGGEEKDE EFQKSESEDSIRRKGFRPSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGAKDV GSENDFADDEHSITEDDIESRRDSLFVPRRHG RNSNLSQTSRSSSMLAAFPANGKMHSTVDC NGVVSLVGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMFKRRSSSFHVSWDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNKGFNMLYIEIKAFVHFM NRYLSYGSGFKRFLVDVLQVALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MIPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVYRLHAV LVHEQQANAGHYWAYIFDIRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETILPPDLRDFV EEDNQRFEKELEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKVLEQPSRSDFSK HLKEETILIQITKASHEHEDKSPETVUGSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKTILLEOFGGRNLSFDERCHNIMKVA QAKLEMIKPEEVNNGLIIMNEFIPFLILLVDENÆE SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIPFPLILLVDENÆE ESGEDREVNNGLIIMNEFIPPFLILLVDENÆE ESGEDREVNNGLIIMNFPIPFLILLVDENÆE	1	{		[
SSGQVAVRNAPQAGSAKAGKGKFQDNFEFIC YFKKFFDANCNEKDYNPVAAGQGQETEVAP SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSJFRLMTQDYWENLYQLTLRA AETTYMIFLVVLVILLGSLYLVTLILAVVAMM YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRKKRKQKEQSGGEKDEI EFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGI RRNSNLSQTSRSSRMLAVFPANGKHISTVDOL NGVVSLVGGSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTINTVE 97 1447 A 1295 2 2057 IQTQLPTKSQQLRGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSBPSQTLPSTTEQQGAL SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPARRHTEEELSVLESCHRWRTEIENDTR DLQESISRHRTIELMYSDKSMQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKVLEQPSRSDFSK HLKEETIQITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENTQRESYDISSLIFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIINNEFIVPFLPLLLVDEMEE ESGEDREVNNGLIINNFIPVPFLPLLLVDEMEE	-06	1446	 	1304		1122	
SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SPDTFSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMIF/LV/LVILLGSLYL/TLILAVWAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAE KLSSKSAKERRNRKKRKQKEQSGGEKDEI EFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGI RRNSNLSQTSRSSMALAVFPANGKMHSTVDC NGYVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSFHVSMDFLEDPSQR QRAMSIASILTNTVE CTSPVDDIDASSPPSGIPSGPGTPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MIPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRHTELMYSDKSMQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQHTKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF GSGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE EGGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE	90	1440	A	1294	1	1456	QLLPPSNRENAGLLVGRCLCSAALRPVGDLIT
SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT SFDTFSWAFLSLFRLMTQDYWENLYQLTLRA AETTYMIFLVVLVILLGSLYLVTLILAVVAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRISDSSSEG KLSSKSAKERNRRKKRKQKEQSGGEKDEI EPQKSESEDSIRRKGFRFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGI RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSFHVSMDFLEDPSQR QRAMSIASILINTVE 97 1447 A 1295 2 2057 QTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNIKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHTEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIOITTKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEVNLEEYEEWHQDVRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMNEFIVPFLPLLLVDEMEE KDILAVEDMRRIRWCSYLGQEMEPHLQEKLT	1	İ					SSGQVAVRNAPQAGSAKAGKGKFQDNFEFIQ
SFDTFSWAFLSLFRIMTQDYWENLYQLTLRA AETTYMIF/LV/LILGSLYLVTLILAV/VAMA YEEQNQAATLEBAEQKEAEFQOMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRKKKKQKEQSGGEKDEL EFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSS PHOSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRHGH RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTINTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRTFYNNNKGPNMLYIEIKAFVHFM NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDVRLHHVVVYFIQNQ APKKIIEKTLLEGFGGRNLSFDERCHINMKVA QAKLEMIKPEEVNLEEYEEWHQDVRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEREVNNGLIMNEFIVPFLPLLLVDEMEE	ľ		i i	i i			YFKKFFDANCNEKDYNPVAAGQGQETEVAP
AETTYMIP/LV/LVILLGSLYLVTLILAV/VAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERNRRKKRKQKEQSGGEKDEI EPQKSESEDSIRRGFRFSIEGNRLTYEKRYSS PHOSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRDSLFVPRHGI RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIGUITKASHEHEDKSPETVLQSAKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIREFEVNLEEYEEWHQDYKKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEREVNNGLIMNSETIVFFIPLLLVDEMEE	Į.	! ;					SIVAPVLNKPNQCPEGYICVKAGRNPNYGYT
AETTYMIP/LV/LVILLGSLYLVTLILAV/VAMA YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATATASEHSREPSAAGRLSDSSSEAS KLSSKSAKERNRRKKRKQKEQSGGEKDEI EPQKSESEDSIRRGFRFSIEGNRLTYEKRYSS PHOSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRDSLFVPRHGI RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIGUITKASHEHEDKSPETVLQSAKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIREFEVNLEEYEEWHQDYKKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEREVNNGLIMNSETIVFFIPLLLVDEMEE							SFDTFSWAFLSLFRLMTQDYWENLYOLTLRA
YEEQNQATLEEAEQKEAEFQQMLEQLKKQQ EAAQQAATAASEHSREPSAAGRLSDSSSEAS KLSSKSAKERRNRKKRÇKÇKEQSGGEEKDEI EPQKSESEDSIRKGFRFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGI RRNSNLSQTSRSSSMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGFKRFPLVDVLQYALEFASSKPV CTSPVDIDIASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKKLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVYFIQNQ APKKIIEKTLLEOFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE	1]	[1 1			AETTYMIF/LV/LVILLGSLYLVTLILAV/VAMA
EAAQQAATATASEHSREPSAAGRISDSSSEAS KLSSKSAKERRNRKKRQKEQSGEKDEIL EFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGF RRNSNLSQTFSRSSRMLAVFPANGKMHSTVDD NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYLHHVVVYFJQNQ APKLIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKPRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE		1					YEEONOATLEEAEOKEAEFOOMLEOLKKOO
KLSSKSAKERNRRKKRKQKEQSGGEEKDEI EFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSS PHQSELLSIRGSLFSPRNSRTSLFSFRGRAKDV GSENDFADDEHSITEDNESRRDSLFVPRRHGI RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATID DNGTTTETEMRKRSSSFHVSMDFLEDPSQR QRAMSIASILTNIVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYSGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHTIEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFPHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE	i	1	İ				EAAOOAATATASEHSREPSAAGRI SDSSSEAS
EFQKSESEDSIRRKGFRFSIEGNRLTYEKRYSS PHQSLLSIRGSLFSPRRNSRTSLFSFRGAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGF RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSSIPSQTLPSTTEQQGALS. SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYMI DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIMNEFIVPFLPLLLVDEMEE		[l [ĺ	KLSSKSAKERRNRRKKRKOKFOSGGEFKDED
PHQSLLSIRGSLFSPRRNSRTSLFSFRGRAKDV GSENDFADDEHSTFEDNESRRDSLFVPRRHGI RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHTEELSVLESCLHRWRTEIENDTR DLQESISRHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLGOPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAYEDMRNRWCSYLGQEMEPHILOEKLT	ł	}					EFOKSESEDSIR DK GEREGIEGNIRI TVEKRVOS
GSENDFADDEHSTFEDNESRRDSLFVPRRHGE RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGGEMEPHLOEKLT)			j j	·		PHOSI I SIDGSI ESDDDNISDTSI ESEDOD AVDV
RRNSNLSQTSRSSRMLAVFPANGKMHSTVDC NGVVSLVGGPSVPTSPVGQLLPEVIIDKPATD DNGTTTETEMRKRRSSSFVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 10TQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEQQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFILCAYQNNIKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGGEMEPHLQEKLT							GSEVIDE A DDEHOTEED VIEW DOOL EVEN DATE
97 1447 A 1295 2 2057 IQTQLPTKSSQQLKGGNCVRCKMQMNFIAE EVLLKYRITFYNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFPHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELL ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT				l l	′		DDNICHT COTCD CCD AT A VIDA NOVA GIOTE DO
DNGTTTETEMRKRRSSSFHVSMDFLEDPSQR QRAMSIASILTNTVE 97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLOEKLT	1						
97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLOEKLT] .						
97 1447 A 1295 2 2057 IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE EVLLKYRITFYNNNKGPNMLYIEIKAFVHFMI NRYLSYGSGPKRPPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTIPSTTEQQGALS SELPSTSPSSVAAISSRSVHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLOEKLT		1			f		
EVLLK YRITFYNNIK GPNML YIEIKAFVHFMI NRYLSYGSGPKRFPLVDVLQY ALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEEELSVLESCLHR WRTEIENDTR DLQESISRIHRTIELMY SDKSMIQVPYRLHAV LVHEGQANAGHY WA YIFDHRESR WMK YNDI AVTKSSWEEL VRDSFGGYRNASAY CLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEOFGDRNLSFDERCHNIMK VA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLOEKLT	-07	1447		1000			
NRYLSYGSGPKRFPLVDVLQYALEFASSKPV CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT	97	144/	Α	1295	2	2057	IQTQLPTKSSQQLRKGGNCVRCKMQMNFIAE
CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT	!					[
SELPSTSPSSVAAISSRSVIHKPFTQSRIPPDLP MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEOFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLQGEMEPHLOEKLT]			.]			NRYLSYGSGPKRFPLVDVLQYALEFASSKPV
MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEOFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLOEKLT	1	1		1			CTSPVDDIDASSPPSGSIPSQTLPSTTEQQGALS
MHPAPRHITEELSVLESCLHRWRTEIENDTR DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQEVDLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEOFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLOEKLT		j		1		Į	SELPSTSPSSVAAISSRSVIHKPFTOSRIPPDLP
DLQESISRIHRTIELMYSDKSMIQVPYRLHAV LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQE\DLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLOEKLT			ĺ	ĺ	ſ	ĺ	MHPAPRHITEEELSVLESCLHRWRTEIENDTR
LVHEGQANAGHYWAYIFDHRESRWMKYNDI AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQE\DLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLOEKLT			ŀ	1	Į.		
AVTKSSWEELVRDSFGGYRNASAYCLMYIN DKAQFLIQE\DLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT		j)	j	ļ]	
DKAQFLIQE\DLIKTGQPLVGIETLPPDLRDFV EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT			i	1		ļ	
EEDNQRFEKELEEWDAQLAQKALQEKLLAS QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT		·	1	j		ļ	
QKLRESETSVTTAQAAGDPKYLEQPSRSDFSK HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT		ļ	J	ļ)	ļ	
HLKEETIQIITKASHEHEDKSPETVLQSAIKLE YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT		, I	Į.				OKI RESETSVITA OA A CODEVI E ODODODODO
YARLVKLAQEDTPPETDYRLHHVVVYFIQNQ APKKIIEKTLLEOFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLOEKLT		ſ	ĺ	[1		HI KEELIUILA SHEREDAGDEMA VOTANTE
APKKIIEKTLLEQFGDRNLSFDERCHNIMKVA QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLOEKLT		l	!		ŀ		VADI VIZI A OPDTEDETENDI I INDIBIDIO
QAKLEMIKPEEVNLEEYEEWHQDYRKFRETT MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT		·		. 1	· 1		
MYLIIGLENFQRESYIDSLLFLICAYQNNKELL SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT					l		
SKGLYRGHDEELISHYRRECLLKLNEQAAELF ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT		i	ì	1	i		
ESGEDREVNNGLIIMNEFIVPFLPLLLVDEMEE KDILAVEDMRNRWCSYLGQEMEPHLQEKLT		}	ł		}	į	MYLIIGLENFQRESYIDSLLFLICAYQNNKELL
KDILAVEDMRNRWCSYLGQEMEPHLOEKLT		j	l	- 1	1		
		- 1		-	[1	ESGEDREVNNGLIMNEFIVPFLPLLLVDEMEE
DELEVIT DESMEIK SEHEPPKI PSYSTHEL CEP		ł	t	1	ĺ	ì	
		ļ	l	i	I		DFLPKLLDCSMEIKSFHEPPKLPSYSTHELCER
FARIMLSLSRTPADGR			l				
98 1448 A 1304 118 453 SGPSSRAIYLHRKEYSONLTSEPTLLOHRVEH	98	1448	A	1304	118	453	
LMTCKQGSQRVQGPEDALQKLFEMDAHGRV						}	
			1	İ	l	l	
WSQDLILQVRDGWLQLLDIETKEELDSYRLD		ľ	- 1		1	4	
99 1449 A 1306 3 1660 CGYFCHTTCAPOAPPCPVPPDLLRTALGVHPE	99	1449	_	1306	3	1660	
COTTCHTTCATQATTCT VITULERIALD VAFE	-	-777	^	1300	•	1000	CGYFCHTTCAPQAPPCPVPPDLLRTALGVHPE
TGTGTAYEGFLSVPRPSGVRRGWQRVFAALS							IGIGIAYEGELSVPKPSGVRRGWQRVFAALS

CEO ID	SEQ ID	1 NAc+	LCEO	Dan dias	I now at the second	LA-L
SEQ ID NO: of	NO: of	Met hod	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	Hou	in NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	Ì	USSN	location	corresponding	r=rnenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		İ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	{	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		{		peptide		/=possible nucleotide deletion, \=possible
L		ĺ		sequence		nucleotide insertion
		Ι			<u> </u>	DSRLLLFDAPDLRLSPPSGALLQVLDLRDPQF
1	1]		ł		SATPVLASDVIHAQSRDLPRIFRVTTSQLAVPP
		l		1		TTCTVLLLAESEGERERWLQVLGELQRLLLD
	1	i		1	1	ARPRPRPVYTLKEAYDNGLPLLPHTLCAAILD
1					1	QDRLALGTEEGLFVIHLRSNDIFQVGECRRVQ
	ĺ		ĺ	ĺ	1	QLTLSPSAGLLVVLCGRGPSVRLFALAELENI
						EVEVPKIPESRGCQVLAAGSILQARTPVLCVA
	1	1	ŀ	I	1	VKRQVLCYQLGPGPGPWQRRIRELQAPATVQ
	1	1			1	SLGLLGDRLCVGAAGGFALYPLLNEAAPLAL
						GAGLVPEELPPSRGGLGEALGAVELSLSEFLL
	1		1	([LFTTAGIYVDGAGRKSRGHELLWPAAPMGW GYAAPYLTVFSENSIDVFDVRRAEWVQTVPL
1			1	1		KK\VRPLNPEGSLFLYGTEKVRLTYLRNQLAE
1						KDEFDIPDLTDNSRRQLFRTKSKRRFFFRVSE
Ì		i				EQQKQQRREMLKDPFVRSKLISPPTNFNHLV
				[HVGPANGRPGARDKSP
100	1450	Α	1318	918	190	SLCVPGPVDTGTFAVMSVMVGSVTESLAPQA
	1	J]]	LNDSMINETARDAARVQVASTLSVLVGLFQV
		1			1	GLGLIHFGFVVTYLSEPLVRGYTTAAAVQVF
		1			į	VSQLKYVFGLHLSSHSGPLSLIYTVLEVCWKL
		1				PQSKVGTVVTAAVAGVVLVVVKLLNDKLQQ
		[QLPMPIPGELLTLIGATGISYGMGLKHRFEAG\
		!				PPVAPNTQLFSKLVGSAFTIAVVGFAIAISLGK
101	1451	A	1353	220 .	445	IFALRHGYRVDSNQVWVMRDV
101	1451	^	2223	220 .	443	DWPDLFTYPLIGSPKCFQSARPE\RMYRRTVR
		j]			SSHGNHALQEVLPRSGHGTEFTKQKHLEAAD HGHPPARMSIFSR
102	1452	A	1363	542	2	AHLLMLNLAL\TDLL\YLTSLPFLIHYYASGEN
				- · -	-	WIFGDFMCKFIRFSFHFNLYSSILFLTCFSIFRY
				j '		CVIIHPMSCFSIHKTRCAVVACAVVWIISLVA
1						VIPMTFLITSTNRTNRSACLDLTSSDELNTIKW
						YNLILTAILLCLPLVIVTLCYTTIIHTLTHGHAN
						\DSCLKQKARRLTILLL
103	1453	Α	1371	2	410	CHSTESSSDFILPGDYLLGGLCPLHSGCLQV\C
	:			ĺ		SFNEHGYHLFQAMRLAVEEINNSTALLPNITL
1		i				GYQLYDVCSDSANVYATLRVLSLPGQHHIEL
	1					QGDLLHYSPTVLAVIGPDSTNRAATTAALLSP
104	1454		1276	2	422	FLVPMLLEQ
104	1434	A	1376	3	432	NSRVEDRS/NMSLWTQNITVCPVRNVTRDGG
t					j	FGPWSPWQPCEHLDGDNSGSCLCRARSCDSP
ì	(!		RPRCGGLDCLGPAIHIANCSRNGAWTPWSSW ALCSTSCGIGFQVRQRSCSNPAPRHGGRICVG
	'					KSREERFCNENTPCPVPIF
105	1455	A	1379	2	396	GLGLLYLIFAAVEGVMRVIGGSNHLAVVLDD
1						IILAVIDSIFVWFIFISLAQTMKTLRLRKNTVKF
				 		SLYRHFKNTLIFAVLASIVFMGWTTKTFRIAK
	1			·		CQSDWMERWVDDAFWSFLF\SLILIVIMFLW
·						RPSA
106	1456	A	1383	1	432	EDGHGGWSSRCLVDHAEEGHREPWKRLCIW
				j		QRGGHEIRFAFYFPGHPLLSPQICLAPETPPRG
					ſ	CPPVSSLHFISLQ/RLPRDCQELFQVGERQSGL
1				ŀ		FEIQPQGSPPFLVNCKMTSGTFWTCRTDSRVF
						QNANPSNAAHSEDQPTP
107	1457	A	1386	719	558	FFFVTRSHSVAQAECSGVFTAHRSLDLVGSSN
100	1.450					YPALSLQSSWDHRHTWLIFAFL
108	1458	A	1397	631	2	RVAISLLCAAIFISFMVQSAGKRWPTGVMLM
l				!		VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT
						VVVLFAFLYSWPIQALLPTYLKTDLAYNPHT VANVLSFSGFGAAVGCCV/GGFLGDWLGTRK AYVCSLLASQLLIIPVFAIGGANVWVLGLLLF

SEQ ID Not of not not of not of not of not of not of not of not of not of not of not of not of not of not of not of not of not of not of not of not not of not not of not not of not not of not not not not not not not not not not	16,
nucle otide sequence uence USSN (09/496) USS	
Contraction Contraction	
Seq- uence	
109	
amino acid residue of peptide sequence T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, **Stop codes Sequence Seq	
residue of peptide sequence	
PQQMLGQGAGILPKLIGGYFDTDQRA	
FTYNVGALGGALAPIIGALIAQRLDLGT	
LSFSLTEVVLRNRRPGKSLVR	GLG
109	ALAS
WSHNSNSMCWGKDQCPYSGCKEALR RVTSRKSAKYRLQGTTPRGDVSLTILNP GVYCRIEVPGWFNDVKINVRLNLQRA RVTSRKSAKYRLQGTTPRGDVSLTILNP GVYCRIEVPGWFNDVKINVRLNLQRA STATE GVYCRIEVPGWFNDVKINVRLNLQRA LAELAFPVGVLATCA*SLLSC*YCVULF FFHSPDALFSLLSCYFSYCFFYLFF CLLLASSPFPLFILLASL FFISTMTKPFEKESEQPA*ATLAFGAQTS CLLKPLSYLNINSSSSSTPATSAGGG TSSSNPPVATFVFGQSSDPVSSYGFVNT SDSLLFSQDSKLATTS GVALKPDLSYLNINSSSSSTPATSAGGG TSSSNPPVATFVFGQSSDPVSSYGFVNT SDSLLFSQDSKLATTS SRRSQRTCSRACSGAWSRTW*RSS*TS STSCSSSSRSCGRGGLGARGVHTSC MYLGEIVRNILIDFTKGFLLRGQISEM MYLGEIVRNILIDFTKKGFLLRGQISEM MYLGEIVRNILIDFTKKGFLLRGQISEM GIFLTFLLSNFLVLLFFVSFVLFQSCL GRKVHLPGHKTGPAVAKDTPEPVKK ATSQG*SFFSEEPPLPPSNEEVPFTLPP* EDP*KNA*LKQMHAATTHWQQHQHQ QYHGIMQ QYHGIMQ QYHGIMQ QYHGIMQ QYHGIMG GRKCGQYWPLEKDSRIRFGFLTVSNLC MNHYKKSTLEILNPEVNPGFFFLTLWKC NYCN THE	
RVTSRKSAKYRLQGTIPRGDVSLTILNP	LYSS
110	DGM
110	ESDS
LAELAFPVGVLATCA*SLLSC*YCVLIFF FFHSPDALFSLLLLSCYPSYCFFYYLFF CLLLASSPFPLFILLASL CLLASSPFPLFILLASL FTSTMTKPFEKESEQPA*ATLAFGAQTS QCALKPDLSYLNNSSSSSSTPATSAGGG TSSSNPPVATIVFGQSSDPVSSYGFVNT SDSLLFSQDSSALATTS TSSNPPVATIVFGQSSDPVSSYGFVNT SDSLLFSQDSALATTS TSSNSPVATIVFGQSSDPVSSYGFVNT SRRSSQRTCSRACSGAWSRTW*RSS*TS STSCSSSSSRSCGRPGGPLGARGVHITSC MSSSTTSSTTSTF HEDIMTHYDRLVDE*ALNAGKQRYEKA MYLGEIVRNILIDFTKGFLLRQJISEM GIFLTFLLSNFLIVCVLLFYVSFYLFQSC ATSQGP*SPFSEEPPLPPSNEEVPPTLPP* EDP*KNA*LKQMHAATTHWQQHQQHQ QYHGIMQ QYHGIMQ QYHGIMQ TIS	STT
	DWM
CLLLASSPPLFILLASL	CSCF
111	SSSPL
112	
TSSSNPPVATFVFGQSSDPVSSYGFVNT. SDSLLFSQDSKLATTS	TAD
112	FGSS
112	LESST
SRRSSQRTCSRACSGAWSRTW*RSS*TS STSCSSSSSRSCGRPGGPLGARGVHITSC MSSSTTSSTTSTF HEDIMTHYDRLVDE*ALNAGKQRYEKN MYLGEIVRNILIDFTKKGFLLRGQISEMI GIFLTFLLSNFLIVCVLLFYVSFYLFQSCI QRTKVHLPGHKTGPAVAKDTPEPVKKE ATSQGP*SFPSEEPPLPPSNEEVPPTLPP* EDP*KNA*LKQMHAATTHWQQHQQHQ QYHGIMQ 115	
STSCSSSSRSCGRPGGPLGARGVHITSC MSSSTTSSTTST	WWR
MSSSTTSSTTSTF	
113	LNSC
MYLGEIVRNILIDFTKKGFLLRGQISEM GIFLTFLLSNFLIVCVLLFYVSFYLFQSCI 114	
114	IISG
114 1464 A 1463 I 396 KQQAVPEPHSSTTTPQEQEQNWYGQDI QRTKVHLPGHKTGPAVAKDTPEPVKKE ATSQGP*SPFSEEPPLPPSNEEVPPTLPP* EDP*KNA*LKQMHAATTHWQQHQQHQQYHGIMQ 115 1465 A 1464 291 2 AGSYPSMVWSCHWGVTQKRRAL*VYS GRRKCGQYWPLEKDSRIRFGFLTVSNLQ MNHYKKSTLEILNPEVNPGFFFLTLWKQ NYCN 116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHN YWTKYQVWEWLQHFLDTNQLDANCIP DINGEHLCSMSLQEFTRAAGTAGQILLYSHLWNGDSLFLCLSLPC 117 1467 A 1479 I 381 GTSGGPKRVLVTERFPWQNPLPVNRGQ VLGPSNSFQRVPLQAQKLVSSHKPGQNG QLQATSVPHPVCMPLNNTQKSKQPLPS/NPEELASDPNNESL*RPWALEDFEIGH KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSÅKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	KTR
114 1464 A 1463 I 396 KQQAVPEPHSSTTTPQEQEQNWYGQDI QRTKVHLPGHKTGPAVAKDTPEPVKKE ATSQGP*SPFSEEPPLPPSNEEVPPTLPP* EDP*KNA*LKQMHAATTHWQQHQQHQQYHGIMQ 115 1465 A 1464 291 2 AGSYPSMVWSCHWGVTQKRRAL*VYS GRRKCGQYWPLEKDSRIRFGFLTVSNLQ MNHYKKSTLEILNPEVNPGFFFLTLWKQ NYCN 116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHN YWTKYQVWEWLQHFLDTNQLDANCIP DINGEHLCSMSLQEFTRAAGTAGQILLYSHLWNGDSLFLCLSLPC 117 1467 A 1479 I 381 GTSGGPKRVLVTERFPWQNPLPVNRGQ VLGPSNSFQRVPLQAQKLVSSHKPGQNG QLQATSVPHPVCMPLNNTQKSKQPLPS/NPEELASDPNNESL*RPWALEDFEIGH KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSÅKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	NFVL
ORTKVHLPGHKTGPAVAKDTPEPVKKE	LNLO
ATSQGP*SPFSEEPPLPPSNEEVPPTLPP* EDP*KNA*LKQMHAATTHWQQHQQHQ QYHGIMQ 115 1465 A 1464 291 2 AGSYPSMVWSCHWGVTQKRRAL*VYS GRRKCGQYWPLEKDSRIFGFLTVSNLG MNHYKKSTLEILNPEVNPGFFFLTLWKG NYCN 116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHN YWTKYQVWEWLQHFLDTNQLDANCIP DINGEHLCSMSLQEFTRAAGTAGQLLYS HLKWNGDSLFLCLSLPC 117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQ VLGPSNSFQRVPLQAQKLVSSHKPGQNI QLQATSVPHPVCMPLNNTQKSKQPLPS/ NPEELASDPNNEESL*RPWALEDFEIGH KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSAKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	FTVP
EDP*KNA*LKQMHAATTHWQQHQQHQQYHGIMQ 115 1465 A 1464 291 2 AGSYPSMVWSCHWGVTQKRRAL*VYS GRRKCGQYWPLEKDSRIRFGFLTVSNLQ MNHYKKSTLEILNPEVNPGFFFLTLWKQ NYCN 116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHN YWTKYQVWEWLQHFLDTNQLDANCIP DINGEHLCSMSLQEFTRAAGTAGQLLYS HLKWNGDSLFLCLSLPC 117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQ VLGPSNSFQRVPLQAQKLVSSHKPGQNI QLQATSVPHPVCMPLNNTQKSKQPLPS/ NPEELASDPNNEESL*RPWALEDFEIGH KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSAKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	POS
QYHGIMQ	VGC
115 1465 A 1464 291 2 AGSYPSMVWSCHWGVTQKRRAL*VYS GRRKCGQYWPLEKDSRIRFGFLTVSNLG MNHYKKSTLEILNPEVNPGFFFLTLWKG NYCN 116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHN YWTKYQVWEWLQHFLDTNQLDANCIP DINGEHLCSMSLQEFTRAAGTAGQLLYSHLKWNGDSLFLCLSLPC 117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQ VLGPSNSFQRVPLQAQKLVSSHKPGQNG QLQATSVPHPVCMPLNNTQKSKQPLPSA NPEEELASDPNNEESL*RPWALEDFEIGH KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSAKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	
GRRKCGQYWPLEKDSRIRFGFLTVSNLO MNHYKKSTLEILNPEVNPGFFFLTLWKO NYCN 116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHN YWTKYQVWEWLQHFLDTNQLDANCIP DINGEHLCSMSLQEFTRAAGTAGQLLYS HLKWNGDSLFLCLSLPC 117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQ VLGPSNSFQRVPLQAQKLVSSHKPGQNG QLQATSVPHPVCMPLNNTQKSKQPLPSA NPEELASDPNNEESL*RPWALEDFEIGH KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSAKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	EEG
MNHYKKSTLEILNPEVNPGFFFLTLWKG NYCN	VEN
116 1466 A 1465 667 337 LPPQRPA*TDSYSTCNVSSGFLAGQSHN YWTKYQVWEWLQHFLDTNQLDANCIP DINGEHLCSMSLQEFTRAAGTAGQLLYS HLKWNGDSLFLCLSLPC 117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQ VLGPSNSFQRVPLQAQKLVSSHKPGQNG QLQATSVPHPVCMPLNNTQKSKQPLPSA NPEEELASDPNNEESL*RPWALEDFEIGH KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSAKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	GEN
YWTKYQVWEWLQHFLDTNQLDANCIP DINGEHLCSMSLQEFTRAAGTAGQLLYS HLKWNGDSLFLCLSLPC 117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQ VLGPSNSFQRVPLQAQKLVSSHKPGQNI QLQATSVPHPVCMPLNNTQKSKQPLPS/ NPEELASDPNNEESL*RPWALEDFEIGH KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSAKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	
YWTKYQVWEWLQHFLDTNQLDANCIP DINGEHLCSMSLQEFTRAAGTAGQLLYS HLKWNGDSLFLCLSLPC 117 1467 A 1479 I 381 GTSGGPKRVLVTERFPWQNPLPVNRGQ VLGPSNSFQRVPLQAQKLVSSHKPGQNI QLQATSVPHPVCMPLNNTQKSKQPLPS, NPEEELASDPNNEESL*RPWALEDFEIGH KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSAKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	HLO
DINGEHLCSMSLQEFTRAAGTAGQLLYS HLKWNGDSLFLCLSLPC 117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQ VLGPSNSFQRVPLQAQKLVSSHKPGQNI QLQATSVPHPVCMPLNNTQKSKQPLPS, NPEEELASDPNNEESL*RPWALEDFEIGE KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSAKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	OEF
HLKWNGDSLFLCLSLPC	NLO
117 1467 A 1479 1 381 GTSGGPKRVLVTERFPWQNPLPVNRGQ VLGPSNSFQRVPLQAQKLVSSHKPGQNG QLQATSVPHPVCMPLNNTQKSKQPLPSA NPEELASDPNNEESL*RPWALEDFEIGH KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSÅKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	
VLGPSNSFQRVPLQAQKLVSSHKPGQNG QLQATSVPHPVCMPLNNTQKSKQPLPS, NPEELASDPNNEESL*RPWALEDFEIGH KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSÅKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	OR
QLQATSVPHPVCMPLNNTQKSKQPLPS/ NPEEELASDPNNEESL*RPWALEDFEIGF KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSAKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	
NPEELASDPNNEESL*RPWALEDFEIGE KGK 118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSAKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	
I 18 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSAKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	PI.G
118 1468 A 1485 3 385 TYLWL*GNPPFYEKNDGGLFELILRAKD PYWDDMSDSAKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	
PYWDDMSDSAKHFIRPLTGRDP*KPFPC QHPWIEGHTCLDNNIHQAASEPINNNFA	EFNS
QHPWIEGHTCLDNNIHQAASEPINNNFA	
H	
119 1469 A 1486 1 398 GTTSKHH*LARSLIRGPFDHDLKPNAAT	יססד
NIIVSYPPTKQLTYEEQDLGWKFRYYLT	
KALTKFLKWVNWDLPQEAKQALELLGI	
PMDVKDSLELLSSHYTNPTVRRYAVARI	DOY
DDEDLLMYL	My/
120 1470 A 1497 3 999 MGESPAV*GYFVLAGMNSAGLSFGGGA	
LAEWMVHGYPSENVWELDLKRFGALO	
FLRHRVMEVMPLMYDLKVPHWDFQTG	
RTSPLYDRLDAQGARWMEKHGFERPKY	
PDKDLLALEQSKTFYKPDWFDIVESEVK	JCK
EAVCVIDMSSFTEFEITSTGDQALEVLQY	LFS
NDLDVPVGHIVHTGMLNEGGGYENDCS	
NKRSFFMISPTDQQVHCWAWLKKHMPF	DZN
LLLEDVTWKYTALNLIGPRAVDVLSELS	IAP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first	Predicted end nucleotide location corresponding to last amino acid residue	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide sequence	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						MTPDHFPSLFCKEMSVGYANGIRVMSMTHT GEPGFMLYIPIEYRWGFTMLSTLVSNS
121	1471	A	1498	3	306	AQFLLVGWDHIL*LIVL*TNLTELGRTTCDQN WPNSPDVLNHGCFYMQCLSKDCTIGYVSRE MLVAHTHTVEEHTGTHLQYVSWPDHSVPDD SSDFVEFEN
122	1472	A	1533	121	329	LGLFSFVWTEVLEEPKDFSCETEDFKTLHCT WDPGTDTALGWSKQPSQSYTLFES*VGSGYII DNFFLA
123	1473	A	1547	111	408	DARTTWKPRNGSSGIWPGDGAK*PPAVEQAE RGHVEMIEKLTFLNLHTSEKDKGGNTALHLA AKHGHSPAVQVLLAQWQDINEMNEKQQTPL HVAADRG
124	1474	A	1555		745	MTFDDDDKNTYGVALVWKKFQTQSLRLSDI. HRKSHLWRGIVSITLIEGRDLKAMDSNGLSDP YVKFRLGHQKYKSKIMPKTLNPQWREQFDF HLYEERGGVIDITAWDKDAGKRDDFIGRCQV DLSALSREQTHKLELQLEEGEGHLVLLVTLT ASATVSISDLSVNSLEDQKEREEILKRYSPLRI FHNLKDVGFLQVKVIRAEGLMAADVTGKSD PFCVVELNNDRLLTHTVYKNLNPEWNKVFTL *VALVWKKFQTQSLRLSDLHRKSHLWRGIVS ITLIEGRDLKAMDSNGLSDPYVKFRLGHQKY KSKIMPKTLNPQWREQFDFHLYEERGGVIDIT AWDKDAGKRDDFIGRCQVDLSALSREQTHK LELQLEEGEGHLVLLVTLTASATVSISDLSVN SLEDQKEREEILKRYSPLRIFNNLKDVGFLQV KVIRAEGLMAADVTGKSDPFCVVELNNDRLL THTVYKNLNPEWNKVFTL
125	1475	A	1556	57	509	GGPAPNSRYAEP*KNSLAMT*AHADCENYVA CGGLDNICSIYNLKTREGNVRVSRELPGHTGY LSCCRFLDDSQIVTSSGDTTCALWDIETAQQT TTFTGHSGDVMSLSLSPDMRTFVSGACDASS KLWDIRDGMCRQSFTGHVSDINAVS
126	1476	A	1592	3	178	KSEKSCVSSLAHFGTSCQRDYDAMVKLVETL EMLPTCDLADQHNIKFHYAFALNR*ER
127	1477	A	1612	1	497	TESPLLVRPYLPYITKSELHAIMTAGFSTIAGS VLGAYISFGVPSSHLLTASVMSAPASLAAAKL FWPETEKPKITLKNAMKMESGDSGNLL*AAT QGASSSISLVANIAVNLIAFLALLSFMNSALA WVGNMFDYPQLSFELICSYIFMPFSFMMGVE WPDSFM
128	1478	A -	1619	286	486	CCMNSKAQESVFKNVLCNPPALSEMPDVKA EDEVDFRASSISEEVAVGSIAATLKMKQGPM TQAINR
129	1479	A	1627	1	395	PTRGALRYWIFGRFLCNIWAAVDVRCCTATI MGLCIISIDRYVGVSYPLRYPTIVTQRRGLMA LLCVWALSLVIYIGPLLGWRHPAPEDETICQI NEEPGYVLFSTPGSFYLPLAIMLVMN*RVYRV AKTE
130	1480	A	1638	2	466	DPRVRTKIVNRKTTIYEIQDKTGSMAVVGKG ECHNIPCEKGDKLRLFCFRLRKRENMSKLMS EMHSFIQIQKNTNQRSHDSRSMALPQEQSQHP KPSEASTTLPESHLKTPQMPPTTPSSSSFTKVT KDKDIK*LLFNLYSSVEILPEVLHLKT
131	1481	A	1651	607	3	LAEGGDVFDCVLNGGPLPESRAKALFRQMVE AIRYCHGCGVAHRDLKCENALLQGFNLKLTD FGFAKVLPKSHRELSQTFCGSTAYAAPEVLQ GIPHDSKKGDVWSMGVVLYVMLCASLPFDD

SEQ ID NO: of NO: of nucl- cotide seq- uence SEQ ID NO: of nucl- cotide seq- uence Seq- uen	idine,
nucl- eotide seq- uence uence Description	•
eotide sequence USSN location corresponding uence USSN location corresponding to last amino acid residue of peptide residue of peptide sequence sequence sequence sequence sequence sequence users are used to location to loc	•
seq- uence 09/496 correspondi ng to first amino acid residue of peptide residue of peptide sequence	line,
uence 914 ng to first arnino acid residue of peptide residue of peptide residue of peptide residue of peptide sequence 914 ng to first arnino acid residue of peptide residue of peptide sequence 92-Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptopl Y=Tyrosine, X=Unknown, *=Stop con /=possible nucleotide deletion, \=possi nucleotide insertion	inie ²
amino acid residue of peptide residue of peptide residue of peptide sequence Peptide sequence residue of peptide sequence Peptide sequence residue of peptide sequence Peptide sequence residue of peptide residue of peptide sequence Peptide residue of peptide re	
residue of peptide /=possible nucleotide deletion, \=possi sequence nucleotide insertion	
peptide /=possible nucleotide deletion, \=possi sequence nucleotide insertion	ini,
sequence nucleotide insertion	hle
	oic .
TDIPKMLWQQQKGVSFPTHLSISA	שוומססו
RLLEPDMILRPSIEEVSWHPWLAS LSNKVGGESKPKKKK	r**KQWQV
132 1482 A 1656 150 48 LVAKSLLYCGCLFFLQLAKNVG	INICENITY A
EANLTSPSPKPTPSSDM*VFLIY*T	VEC V MUTA
VDAO	II OAWIIV
133 1483 A 1660 3 406 RKHIKLLIQKLSDVP*ECQNNQL*1	LTEICEKE
KKEFKKKMDDQRPEKITEA*SKDI	SPMEEEK
TEMIRSYIQEVGRYIKRLEEAQSKI	RLEKLREK
HKERQPILDEKPKGEGSSSFLSET	CHEDTSWF
PNFTP	
134 1484 A 1666 1276 466 PGSTHASARITIY*L*IILSNATEVD	NNFSKPPP
FFPAGAPPASSSSSSSSSPPTVSTA	
PPPPGAPPPSLIPTIESGHSSGYDSR	
NVAFPHLPGSAPSWPSLVDTSKQV	
SSSSSSSSSSPRDRDRER*RTRE	RERERDHS
PTPSVFNSDEERYRYREYAERGYE	RHRASRE
KEERHRERRHREKEETRHKSSRSN	SRRRHESE
EGDSHRRHKHKKSKRSKEGKEAG	SEPAPEOE
STEATPAE	
135 1485 A 1673 1 417 PTRPVNSSQAFALVYYTLGALGGN	ILIAHMGI.
GYRYWAGIGVLQSCESALTHYRL	VANHVAS
DISLTGGSVVQRIRLPDEVENPGMI	NSGMLOE
DLIQYYQFLAEKGDVQAQVGLGQ	
GV*QNHQRAFDYFNLAA	
136 1486 A 1678 525 9 ANTSLSSAAVSAVSPPPCRTSTATT	LPPPMPSF
FCVFPSPSMSPSPSEFLSCIASVSRV	
GSSSTASSLNFSAIMGSSSATASWV	LSTASTPP
CPSALPSSPAQES*SLAASSSAWPV	AGISPSGA
CTFPAGSASGAAKAPSPSWRCPSFI	RALFSLLD
SSSLSL	
137 1487 A 1680 1 2999 AHRDEIQRKFDALRNSCTVITDLEE	QLNQLTE
DNAELNNQNFYLSKQLDEASGAN	DEIVQLRS
EVDHLRREITEREMQLTSQKQTME	ALKTTCT
MLEEQVMDLEALNDELLEKERQW	EAWRSVL
GDEKSQFECRVRELQRMLDTEKQS	RARADQ
RITESRQVVELAVKEHKAEILALQC	ALKEQK
LKAESLSDKLNDLEKKHAMLEMN	ARSLQQK
. LETERELKQRLLEEQAKLQQQMDI	QKNHIFR
LTQGLQEALDRADLLKTERSDLEY	QLENIQV
LYSHEKVKMEGTISQQTKLIDFLQA	KMDQPA
KKKKVPLQYNELKLALEKEKARCA	
KTRIELRSAREEAAHRKATDHPHPS	
QIAMSAIVRSPEHQPSAMSLLAPPS	KRKESST
PEEFSRRLKERMHHNIPHRFNVGLM	MRATKC
AVCLDTVHFGRQASKCLECQVMC	#KCSTC
LPATCGLPAEYVTHFTEAFCRDKM	
KEPSSSLHLEGWMKVPRNNKRGQO	ZGWDRK
YTVLEGSKVLIYDNEAREAGQRPVE	
DGDVSIHGAVGASELANTAKADVI	
HPHTTCWPGRTLYLLAPSFPDKQR	
VVAGGRVSREKAEADAKLLGNSLI	KLEGDD
RLDMNCTLPFSDQVVLVGTEEGLY	ALNVLK
NSLTHVPGIGAVFQIYIIKDLEKLLM	IIAGEERA
LCLVDVKKVKQSLAQSHLPAQPDI:	
KGCHLFGAGKIENGLCICAAMPSK	
ENLSKYCIRKEIETSEPCSCIHFTNY	
FYEIDMKQYTLEEFLDKNDHSLAPA	VFAASS
NSTPVSIVQVNSAGQREEYLLCFHE	FGVFVDS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
ectide	seq-	J	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		}	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
j,	1	j	j	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ļ	l		peptide	sequence	/=possible nucleotide deletion, \=possible
ļ]	1	l	sequence	1	nucleotide insertion
		 		buqueine	 	YGRRSRTDDLKWSRLPLAFAYREPYLFVTHF
						NSLEVIEIQARSSAGTPARAYLDIPNPRYLGPA
ł		1				ISSGAIYLASSYQDKLRVICCKGNLVKESGTE
	-	1	1		!	HHRGPSTSRR*PASPLPQYQGQRAFLQGRRK
138	1488	A	1686	2	526	GRPQGPAPGAGSPPESGPGLWAALGCSLVWV
1				_	} ===	PLCCLGGAAGRL*ARSGKSGLRRRRAHAGPP
						PGGPCNSCP*CSAPESGGRGPLPGPGTGGVCS
1 1		ł			ł	CWTRGCQTTARTAAAAAAPGPAGRRPPGGA
					•	PQNGSCAASASQEAAAPPPMCPPGRRWAVAS
1		ł	ł .		ł	PPETRCPAAPGTRCRRLEAA
139	1489	A	1693	3	376	LPSMSNCTSCFRLQSRTES*IRQAGHLLGRNE
					1 5.0	FIETKALGCAWFSLCYYLVLYFESSHKVDFVF
1 1					i	IV*CFSTPPGAQMTIMSQACAERCNIMRLVDR
l Ì		ļ				RWAGIAKGVGTQKIIGRVHLGEQKALGL
140	1490	A	1704	3	376	ERTNKFIKELIMDGKNLIAATKSLSVAQRKFA
	,	•	•,, • .		[370	HSLRDFKFEFIGDAVTDDERCIDASLREFSNFL
! !						KNLEEQREIMVS*EGCKLISQLSRGKKIWIWK
						LVLVEVVKHLSLGTVVHCNGKMRFPEP
141	1491	A	1743	1	362	LITNKVFVARELSCLDVHLDSTGSTAVVADQ
			1, 15	•	302	DKLELELVLKGSYEDTQTSFLGTASAFRFHY
i i						MAAL*TELSGRLRSSKSNGWNGDNSTGYLTV
ĺĺ						PLRPLTIVKEVTMDVPAPNVRGLNWMG
142	1492	A	1769	1	406	NNPSTLPRGS*PMSPRTTMGRRRQRRREHKSS
			•,•,•	•	100	LSLASSTVGPGGQIVHTETTEVVLCGDPLSGF
						GLQLQGGIFATETLSSPPLVCFIEPDSPAERCG
1					1	LLQVGDRVLSINGIATEDGTMEEANQLLRDA
					•	ALAHKVV
143	1493	Λ	1789	1	447	QMLRNGGDQNTVPDYHFADRIRELL*PTEDQ
1						KNCIP*DTYLRPSALGNIVEEVTHPCSPGPCPA
						NELCEVNRKGCTSGDPCLPYFCVQGCKLGQA
}						SDFIARQGTLIQVPSSAGEVECYKICSCGQSGL
			1			LENCMEMHCMDLPTDTSALVR
144	1494	Α	1814	1	404	PGRRFRPRLSQAGTDSGS*VFPDSFPSAPAEPL
1		ĺ		į		PYFLQEPQDAYIVKNKPVELRCRAFPATQIYF
	1		j	ļ	j	KCNGEWVSQNDHVTQEGLDEATGLRVREVH
						IEVSRQQVEELFGLEDYWCQCVAWSSAGTTK
l						SRRAYVRI
145	1495	A	1827	26	448	XVEEKHADTWRSXCLSDFFFHAAKXLCXE*N
1		- 1	İ		ł	CGDAISLSVGDHFGKGNGLTWAEKFQCEGSE
	Į	ļ	I			THLALCPIVQHPEDTCIHSREVGVVCSRYTDV
İ		ĺ	İ	i		RLVNGKSQCDGQVEINVLGHWGSLCDTHWD
						PEDARVLCRQLNCGTAL
146	1496	A	1828	574	333	QHEGGDLRRRQLGEIQLTVRYVCLRAASAC*
	ĺ	l				SMAAET*HHVPASGADPYVRVYLLPERKWA
				1		CRKKTSVKRKTLEPLFDET
147	1497	A	1855	1	372	ERLVLTSEHCLVLTLFWPSWTYHTLLLSRQH
	ŀ	- 1	1	ł	j	VRRLPKLTHAEHDHLASIMNKLLTNYDNLFE
	ļ			l		TSVTYSMG*HGAPTGSEAGANWNH**LHAH
	1	1	- 1	1		YYPPLLRSDTVRKFMVGSQMLAQAQRDLTPE
						Q
148	1498	A	1879	568	7	LLSALDDKGGTQPSASFSNAPTIVCVTACPAG
	l	[,		IAHTYMAAEYLEKAGRKLGVNVYVEKQGAN
į	1	Į	ĺ	j	ļ	GIEGRLTADQLNSATACIFAAEVAIKESERFN
J	J	j	ŀ	l		GIPALSVPVAEPIRHAEALMQQALTLKRSDET
	1		1	[ĺ	RTVQQDTQPVKSVKTELKQALLSGISFAVPLI
						VAGGTQVA*AV*RQGISSLHDVQVRTWNS
149	1499	A	1880	611	24	GLNSENALSNEAMERGWQCLRLFAERLQDIP
Į	j		1	!	ł	PSQIRVVATATLRLAVNAGDFIAKAQEILGCP
						VQVISGEEEARLIYQGVAHTTGGADQRLVVD
						

CEC TO	CEOTIN	Met	SEQ	Dendistad	Dundistad and	Amino said reguence (AmAlasias Come)
SEQ ID NO: of	SEQ ID NO: of	hod	ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	}	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
1	İ		***	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ]	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ŀ			peptide	1 1	/=possible nucleotide deletion, \=possible
j		ļ	}	sequence		nucleotide insertion
						IGGASTELVTGTGAQTT*LFSLSMGCVTWLER
						YFADRNLGQENFDAAQKAAREVLRPVADEL
		l				RYHSWKEVRGASVTVQALQEIMMAQGMDE
	Ĺ	<u> </u>		ĺ <u> </u>		RITMEIWPVD
150	1500	A	1894	2	750	GRVDFFHTDYRPLIRDSNNYVLDEQTQQAPH
		1			İ	LMPPPFLVDVDGNPHPTKYQRLVPGRENSAD
		Į				EHLIPQLGYVATSDGEVIEQIISLQTNDNDERS
		İ	Ĭ			PESSILDGMIRQLQQQQDQRMGADQDTIPRG
					İ	LSNGEETPRRGFRRLSLDIQSPPNIGLRRSGQV
		ĺ	1			EGVRQMHQNAPRSQIATERDLQAWKRRVVV
	ĺ	ŀ				PEVPLGIFRKLEDFRLEKGEEERNLYIIGRKRK
L.,.		<u> </u>				TLQLSHKSDSVGLVSQSRPRTCRRKYP
151	1501	A	1900	141	785	GKTIQIQTTMQNKYKTVQKQYKTIPKNKKA
						MEMQIKKQFQDTCKVQTKQYKALKNHQLEV
		l	1			TPKNEHKTILKTLKDEQTRKLAILAEQYEQSI
						NEMMASQALRLDEAQEAECQALRLQLQQEM
						ELLNAYQSKIKMQTEAQHERELQKLEQRVSL
			1			RRAHLEQKIEEELAALQKERSERIKNLLERQE
152	1502	Ā	1915	2	377	REIETFDMESLRMGFGNLVTLDFPKEDYR
132	1302	^	1913		311	LVRLLDTÖRDGLÖNYEALLGLTNLSGRSDKL ROKIFKERALPDIENYMFENHDOLROAATEC
		1				MCNMVLHKEVQERFLADGNDRLKLVVLLCG
						EDDDKVQNAAAGALAMLTAAHKKLCLKMT
						OVTT
153	1503	A	1921	1	237	AYQSLRLEYLQIPPVSRAYTTACVLTSAAVQL
	1505	''	.,	•	23,	ELITPFQLYFIPELIFKHFQIWRLITNFLFFVPFG
]			FNFLLYMIFLYT
154	1504	A	1928	2	354	EMVEGGEGKMCINTEWGGFGDNGCIDDIRTR
	,					YDTEVDEGSLNPGKQRYEKMTSGMYLGEIV
						RQILIDLTKQGLLFRGQISERLRTRGIFETKFLS
		}	1			QIESDRLALLQVRRILQQLGLD
155	1505	A	1929	2	369	TEIAKIKMEAKKKYEKELTMFQNDFEKACQA
						KSEALVLREKSTLERIHKHQEIETKEIYAQRQ
						LLLKDMDLLRGREAELKQRVEAFESYQLELK
						DDYIIRTYRLIEDDRINIQISGHWQESP
156	1506	Ä	1935	1	270	VTRKLPIFIVDAFTARAFRGSPAADCLLENEL
	!					DEDMHQKIAREMNLSETAFIRKLHPTDNFAQ
	,,,,,					RSCFGLIWFTPTTDLQILTSSILPSIL
157	1507	A	1936	584	305	ESKVNNEKFRTKSPKPAESPQSATKQLDQPTA
	•					AYEYYDAGNHWCKDCNTICGTMFDFFTHMH
			السييا			NKKHTQGQFQKSSDFQKEELQQTFLPPERQG
158	1508	Α	1939	1	423	TTHRLNVTAEPPCTSMPIYWMPDVPHRCTTA
	*			·		NTCPVDLTDYCAQNGFYCLVYGFLPYGSLED
						RLHCQTQACPPLSWPQRLDILLGTARAIQFLH
						QDSPSLIHGDIKSSNVLLDERLTPKLGDFGLA
150	1600		1074		40.	RFSRFAGSSPIQSSM
159	1509	A	1974	3	401	HTSTARLLLHRGAGKEAVTSDGYTALHLAAR
						NGHLATVKLLVEEKADVLARGPLNQTALHL
				. ,	-	AAAHGHSEVVEELVSADVIDLFDEQGLSALH
						LAAQGRHAQTVETLLRHGAHINLQSLKFQGG
160	1610		1000	2	415	HGPAATLLR
160	1510	A	1982	2	417	KFLKDLEKQYNKEEPHLSEIGSCFLQNQEGFA
l						IYSEYCNNHPGACLELANLMKQGKYRHFFEA
						CRLLQQMIDIAIDGFLLTPVQKICKYPLQLAEL
						LKYTTQEHGDYSNIKAAYEAMKNVACLINER
161	1511	A	1984	4	770	KRKLESIDKIA
101	1211	Λ.	1704	*	110	RETGSVSLSPSGLEGAESYAVSPILYSSPDVKE
						LWLETLQGQRHSHTGVKSTPGQSAAILMKLR
						SSHNASKTLNANNMETLIECQSEGDIKEHPLL

SEO ID	SEQ ID	Met	SEQ	Predicted	Dendicted and	L Amino ocid popusaco (A=A), ' - C C
NO: of	NO: of	hod	ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		{	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	Î	ĺ	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	}	1	peptide	50425200	/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
			 			ASCESEDSICQLIEVKKRKKVLSWPFLMRRLS
-			i i			PASDFSGALETDLKASLFDQPLSIICGDSDTLP
j	j		}	}		RPIQDILTILCLKGPSTEGIFRRAANEKARKEL
			1			KEELNSGDAVDLERLPVHLLAVVFKDFLRSIP
L		ł	1		İ	RKLLSSDLFEEWMGALEMQDEEDRIEALK
162	1512	A	1986	864	501	LLNSGLFSAPDGSNLEMRLTRGGNMCSGRIEI
			1			KFQGRWGTVCDDNFNIDHASVICROLECGSA
1	i	ł	i	ł	ì	VSFSGSSNFGEGSGPIWFDDLICNGNESALWN
		<u> </u>			<u> </u>	CKHQGWGKHNCDHAEDAGVICSSKD
163	1513	Α	2001	419	187	AVDLSIDESSLTGETTPCSKVTAPOPAATNGD
	ľ	Ì				LASRSNIAFMGTLVRCGKAKGVVIGTGENSE
L	<u> </u>	<u> </u>				FGDIINLSTFVVHS
164	1514	A	2012	284	597	SLLCLFPGTSTVVCKPIVIETQLYVIVAQLFGG
1	Ì				1	SHIYKRDSFANKFIKIQAIEILKIRKPNDIETFKI
ļ) ·	ļ	1 1			ENNWYFVVADSSKAGFTTIYKWERETGFYSH
1/2	<u> </u>					QSFTR
165	1515	Α	2013	2	403	EDPBELGHFYDYPMALFSTFELFLTIIDGPANY
ĺ		l	1 1			NVDLPFMYSITYAAFAIIATLLMLNLLIAMMG
		ĺ	1 1			DTHWRVAHERDELWRAQIVATTVMLERKLP
ĺ						RCLWPRSGICGREYGLGDRWILRVEDRQDLN
166	1516	A	2019	2	000	RQRIQRYA
100	1210	Α	2019	2	927	CCQREGLGLKAVVQILLSHGRNGLPGEPASS
1	-		!			QGLSAASSTPVFHLALQIDSAPDNIDWVEMLF
}		l				NKNMVTERLQNVMVLEQCFSDSSSLYRFLTY
}			1	. (SYLLAFNVWLLLAPVTLCYDWQVGSIPLVETI
						WDMRNLATIFLAVVMALLSLHCLAAFKRLE HKEVLVGLLFLVFPFIPASNLFFRVGFVVAER
l	}			'		VLYMPSMGYCILFVHGLSKLCTWLNRCGATT
		1		ĺ		LIVSTVLLLLLFSWKTVKQNEIWLSRESLFRS
				.		GVQTLPHNAKVHYNYANFLKDQGRNKEAIY
			1		[HYRTALNNNKAWDYLCWRFRKTLTDLP
167	1517	A	2025	696	71	AAASAASSLTVTLGRLASACSHSILRPSGPGA
	[[[ASLWSASRRFNSQSTSYLPGYVPKTSLSSPPW
1						PEVVLPDPVEETRHHAEVVKKVNEMIVTGQY
			1 1	1		GRLFAVVHFASRQWKVTSEDLILIGNELDLA
				i		CGERIRLEKVLLVGADNFTLLGKPLLGKDLV
					ļ	RVEATVIEKTESWPRIIMRFRKRKNFKKKRIV
160			<u> </u>			TTPQTVLRINSIEIAPCLL
168	1518	A	2046	2	366	HLQVAARVFMPLQAVDSAPKPLKGQAQAPQ
		i		1		RLQGAARVFMPLQAQVKAKASKPLQMQIKA
				l		PPRLRRAARVLMPLQAQVRAPRLLQVQSQVS
1/0	1,516		1			KKQQAQTQTSEPQDLDQVPEEFQGQDQVLR
169	1519	A	2049	1	945	QNLEDREVLNGVQTELLTSPRTKDTLSDMTR
				l		TVEISGEGGPLGIHVVPFFSSLSGRILGLFIRGI
	.[[I		EDNSRSKREGLFHENECIVKINNVDLVDKTFA
				j	Ì	QAQDVFRQAMKSPSVLLHVLPPQNREQYEKS
	l	ĺ		ł	ł	VIGSLNIFGNNDGVLKTKVPPPVHGKSGLKTA
	ŀ			i	1	NLTGTDSPETDASASLQQNKSPRVPRLGGKPS
			}	ļ	ì	SPSLSPLMGFGSNKNAKKIKIDLKKGPEGLGF
]	ļ			ĺ	TVVTRDSSIHGPGPIFVKNILPKGAAIKDGRLQ
	- 1		j	ļ	j	SGDRILEVNGRDVTGRTQEELVAMLRSTKQG
170	1520	Ā	2050	363	1	ETASLVIARQEGHFLPRELVMFRSQSH
***	1320	^	2030	202		PVATHLTKILNSDEHAVVISSAKTLCETVKDF
İ	Į	[{	ĺ	VAKVEKTYDKTLENAVVADAVASKCSVLNE
	l			ŀ	l	KLEQLLQALHTDSQAAPVLPGLSPLIVEEDAV
171	1521	A	2055	139	675	ESSSEESLGESKEQLGDDVTKPSSQKA
··-		- · ·	2000		- [IPSRPWLGRITGLDPAGPLFNGKPHQDRLDPS DAGEVDVIHSDTDALGYKEPLGNIDEVENGG
ļ			1	j	ŀ	DAQFVDVIHSDTDALGYKEPLGNIDFYPNGG
						LDQPGCPKTILGGFQYFKCDHQRSVYLYLSSL

NO: of No: of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
							D=Aspartic Acid F=Glutomic Acid
Sequence				1 .			
1949 09496 007490 0074	1	,	1				
			ŀ				
mino acid residue of peptide sequence T-Threanine, W-Valine, W-Tryptophan, Y-Stop codon, peptide sequence T-Threanine, W-Valine, W-Tryptophan, Y-Stop codon, peptide sequence T-Threanine, W-Valine, W-Tryptophan, Y-Stop codon, peptide sequence T-Threanine, W-Valine, W-Tryptophan, Y-Stop codon, peptide sequence T-Threanine, W-Valine, W-Tryptophan, Y-Stop codon, peptide sequence T-Threanine, W-Valine, W-Tryptophan, Y-Stop codon, peptide sequence T-Threanine, W-Valine, W-Tryptophan, Y-Stop codon, peptide sequence T-Threanine, W-Valine, W-Tryptophan, Y-Tryptophan, Y-Stop codon, peptide sequence T-Threanine, W-Valine, W-Tryptophan, Y-Trypt			ł				O-Glutomine D-Assising C-Cosi-
Persidue of popular			1	1			TeThreenine VeVeline WeTender
	ŀ		l				
			ļ	1		sequence	/
RESCTITA-YPCDSYQDYRNGKCVSCOTISQUE	1						nucleotide insertion
SCPLLGYYADNWKDHLRGKDPPMKAFEDT			 		sequence	 	DESCRIPTA VIOLOGIA OPTIONALISMO COMPONENTI
172	1		İ	[1	COLL CVV A DARREDON DON CONTRACTOR
172			j				A CECUTO A CONTROL TO THE CONTROL TH
MFWKVKIRSOTRYTTREPVDRILKERALU KERRSGITTDDDTMSEMKMGRYWSKEERKO H-VROKEORRREPMMRILKCLKES 173	172	1522		2056	3	261	
	1/2	1322	Λ	2030		301	LIQHKSAVEYAQSHLSLVSMCKESHKCSEPK
HLVRGKEGRRREFMMRIKLKCLKES			1			İ	MEWKVKIKSDGIRYIIKKPVKDRILKERALKI
173	1		i	1		ł	KEEKSGLI I DDDI I MSEMKMGRY WSKEERKQ
	172	1500		200		205	HLVRGKEQRRRREFMMRIRLKCLKES
AVFLSVIVILTYTGYILPWSGRFYSLWDTIGYA	1/3	1525) A	2000	1	387	
174							
174	1 1			1		j	
174	1 1				· ·		
RRLKEBEEARLKYEKEEMERLEIQRIEKEKW RRLEAKDLERKNEELEELYLLERCFPEAELKK QETKLLSQWKHYIQCDEPSPSVAQEMNT	I						
HRLEAKDLERRNEELEEL/ILLERCFPEAEKLK	174	1524	A	2071	74	443	LLMGPKAKKSGSKKKKVTKAERLKLLQEEEE
175	1 1	,					
175	i i		1				
175	L	1000	<u> </u>	السيا			QETKLLSQWKHYIQCDGSPDPSVAQEMNT
RRVLGGLPLPSPAPMPIMSLPEGESRKEREVQ RLQPFYLEPGHELPATTLLAFLAAV 176	175	1525	A	2083	139	486	AALTWSQPQEFWPMEMQPIVTDMVTVHWV
RI.QFPYLEPGHELPATTILAFLAAV	1	'		[[AESSTVGWLCALFRVTHVGVGATGHGVVCG
1526	İ			l l			RRVLCGLPLPSPAPMPIMSLPEGESRKEREVQ
FQKFLNLLGDTITLKGWTGYRGGLDTRNDTT GHSVYTVYQGHEMFHVSTMLPYSKENKQQ VEKKRHIGNDIVTIVFQSGESSPAFKPSMRS HFTHIFALVRYNQQNDNYRLKIFSEESVPLFG PPLPTPPVFTDHQEFRDFLLVKLINGEKATLET PCI							RLQFPYLEPGHELPATTLLAFLAAV
FOKFENILLGDTITLKGWTGYRGGLDTKNDTT GHSVYTVYOGHEIMFIVSTMLPYSKENKQQ VERKRHIGNDIVTIVFQEGESSPAFKPSMIRS HFTHIFALVRYNQQNDNYRLKIFSEESVPLFG PPLPTPPFTDHQEFRDFLLVKLINGEKATLET PCI	176	1526	Α	2092	3	587	EGSVNFKFGVLFAKDGQLTDDEMFSNEIGSEP
]						FQKFLNLLGDTITLKGWTGYRGGLDTKNDTT
VERKRHIGNDIVTIVPQEGESSPAFKPSMIRS	!!!						GIHSVYTVYQGHEIMFHVSTMLPYSKENKOO
HFTHIFALVRYNQQNDNYRLKIFSEESVPLFG PPLPTPPVFTDHQEFRDFLLVKLINGEKATLET PPC	1 1						
PPLPTPVFTDHQEFRDFLLVKLINGEKATLET]]			ļ į			
PCI	[1						PPLPTPPVFTDHQEFRDFLLVKLINGEKATLET
CDGAWLAWACWVFGNDFPSFASAACSALLG CSVSTACLCVPLCSGSPLAPFRTAALQEGLR RAVSVPLTLAETVASLWPALQELARCGNLAC RSDLQ PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVPPNIQNPD 179 1529 A 2111 1 312 PIRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQITEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVCTFHLTACVSLHRIHNSTVV QAFASVCTFHLTACVSLHRIHNSTVV SUGGAHIERLFQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWYTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRNPQSWFDN GSICI BS2 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWYLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	<u> </u>						
CSVSTACLCVPLCSGSPLAPFRRTAALQEGLR RAVSVPLTLAETVASLWPALQELARCGNLAC RSDLQ 178 1528 A 2104 2 409 ALQSTLGAVWLGLLLNSLWKVAESKDQVFQ PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD 179 1529 A 2111 1 312 PIRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFERQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRHINSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRNPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	177	1527	A	2103	44	427	GKGQVSLEGRPHRGPLCLGSWWPGSRVPGC
CSVSTACLCVPLCSGSPLAPFRRTAALQEGLR RAVSVPLTLAETVASLWPALQELARCGNLAC RSDLQ 178 1528 A 2104 2 409 ALQSTLGAVWLGLLLNSLWKVAESKDQVFQ PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD 179 1529 A 2111 1 312 PIRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFERQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRHINSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRNPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF				1	, •		
RAVSVPLTLAETVASLWPALQELARCONLAC RSDLQ 178 1528 A 2104 2 409 ALQSTLGAVWLGLLLNSLWKVAESKDQVFQ PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVPFPIQNPD 179 1529 A 2111 1 312 PIRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRHINSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRIPPQSWFDN GSICI 182 1532 A 2123 I 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF		i		1			CSVSTACLCVPLCSGSPLAPFRRTAALOEGLR
RSDLQ	1	i	1				RAVSVPLTLAETVASLWPALOELARCGNLAC
178 1528 A 2104 2 409 ALQSTLGAVWLGLLLNSLWKVAESKDQVFQ PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD 179 1529 A 2111 1 312 PIRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLEGGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV GAFASVVCTFHLTACVSLHRIHNSTVV GGAFASVVCTFHLTACVSLHRIHNSTVV GGAFASVVCTFHLTACVSLHRIHNSTVV GGAFASVVCTFHLTACVSLHRIHNSTVV GGNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI TSIS A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLIVVILENIAVCF	1 1						RSDLO
PSTAASSEGAVVEIFCNHSVSNAYNFFWYLHF PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD 179 1529 A 2111 1 312 PIRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAHILFEREQEHLAPGGISMPL PPPLPLACLG LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWYTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	178	1528	Α	2104	2	409	
PGCAPRLLVKGSKPSQQGRYNMTYERFSSSL LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386. YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF						•	PSTA ASSEGA V VEIECNHS V SNA V NEEWYL HE
LILQVREADAAVYYCAVEVPNTDKLIFGTGT RLQVFPNIQNPD 1529 A 2111 1 312 PIRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVWFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 I 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	ĺ	ł					PGCAPRLLVKGSKPSOOGRYNMTVERESSSI
RLQVFPNIQNPD 1529							L.II.OVREADAAVYYCAVEVPNTDKI IFGTGT
179 1529 A 2111 1 312 PTRSSTRPPSLFVHASAKGGEKEEGDDGHYL MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 I 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	l 1			}			
MRTESHTGLKKGGNANLVFMLKRNTEPKKG SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG 180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV GAFASVVCTFHLTACVSLHRIHNSTVV STEPPPLYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWYTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	179	1529	A	2111	1	312	
SYHFDLERLRAAHILFEREQEHLAPGGISMPL PPPLPLPACLG		1			-		
180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 I 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF		1			1		SYHEDI ERI RAAHII FERFOFUI ADCCIELADI
180 1530 A 2116 3 366 TSIKRAIETTDVTRSFGWDSSEAWQQHDVQE LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF							PPPLPLPACIG
LCRVMFDALEQKWKQTEQADLINELYQGKL KDYVRSLECGYEGWRIDTYLDIPLVIRPYGSS QAFASVVCTFHLTACVSLHRIHNSTVV 181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	180	1530	A	2116	3	366	
181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 I 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF		-000	1	2110	-	200	
181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 I 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF		1					KUANDEL ECCAECINDIDANI DIDI ADDIACO
181 1531 A 2117 2 386 YGLGAHFGRLFIQAGINENDFYDGAWCAGR NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 I 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	[. [[(.	j		
NDLQQWIEVDARRLTRFTGVITQGRNSLWLS DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	181	1531		2117		206	
DWVTSYKVMVSNDSHTWVTGKNGSGDMIFE GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 I 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMOFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	101	1221	Λ	211/	- .	200.	
GNSEKEIPVLNELPVPMVARYIRINPQSWFDN GSICI 182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGRPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF		1	1	ŀ	ļ	ļ	
182 1532 A 2123 1 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF			1]	[1	
182 1532 A 2123 I 493 RTKTDVYILNLAVADLLLLFTLPFWAVNAVH GWVLGKIMCKITSALYTLNFVSGMQFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF		ļ	l	l	ł	ł	·
GWVLGKIMCKITSALYTLNFVSGMOFLACISI DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	100	1530					
DRYVAVTKVPSQSGVGKPCWIICFCVWMAAI LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	102	1052	Α	2123	1 /	493	
LLSIPQLVFYTVNDNARCIPIFPRYLGTSMKAL IQMLEICIGFVVPFLIMGVCYFITARTLMKMP NIKIS 183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	1	ľ		ļ	l		
183 1533 A 2140 3 S61 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF]			1		
183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF			1		1	İ	•
183 1533 A 2140 3 561 RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	i	- 1	i		1	ľ	IQMLEICIGFVVPFLIMGVCYFITARTLMKMP
RSGNYFVTMFDDYSATLPLLIVVILENIAVCF		l	l				
RSGNYFVTMFDDYSATLPLLIVVILENIAVCF	183	1533	A	2140	3	561	RQAWHEAFKVRKEILTVICCLLAFCIGLIFVQ
		1	I		I		
							VYGIDKFMEDLKDMLGFAPSRYYYYMWKYI

SEQ II	D SEQ ID	Met	SEQ	Predicted	Predicted end	
NO: of		hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	mod	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		i i	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
l	1		ł	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	·	1	1	residue of	sequence	Y=Tyrosinc, X=Unknown, *=Stop codon.
	Ì			peptide		/=possible nucleotide deletion. \=possible
		L	ļ	sequence		nucleotide insertion
ł	1	ł				SPLMLLSLLIASVVNMGLSPPGYNAWIEDKAS
	ł	i			1	EEFLSYPTWGLAVCASLDVFAILPVPVAFIGR
	1.534	 	10115	 		RFSLIDDGAGPFCSAAYTTTGCRTPYL
184	1534	Α	2145	3	538	HELTVAAADRGQPPQSSVVPVTVTVLDVND
ļ.		1				NPPVFTRASYRVTVPEDTPVGAELLHVEASD
	{	1		ļ	ļ	ADPGPHGLVRFTVSSGDPSGLFELDESSGTLR LAHALDCETQARHQLVVQAADPAGAHFALA
[1	ſ		[ĺ	PYTIEVQDVNDHGPAFPLNLLSTSVAENQPPG
	1			Ì		TLVTTLHAIDGDAGAFGRLRYHL
185	1535	A	2151	2	671	LDKLLDRMENYNIFNEYILKQVAATYIKLGW
	1333	1	1	[-	***	PKNNFNGSLVQASYQHEELRREVIMLACSFG
l	- [1		1		NKHCHQQASTLISDWISSNRNRIPLNVRDIVY
Į	į		1	!	ì	CTGVSLLDEDVWEFTWMKFHSTTAVSEKKIL
	į.		İ			LEALTCSDDRNLLNRLLNLSLNSEVVLDQDAI
ĺ		1	ĺ .	ĺ	ľ	DVIIHVARNPHGRDLAWKFFRDKWKILNTRI
						RQKTLEFDFAEPLILAFPIILYTAIDNPPLVREH
						E
186	1536	A	2153	2	400	GPMCDKHSAFAEKFHAGFIDYTVHPLWETWA
						HLALPDAQDILYTLEDNRNWVDSMIPQSPSPP
]		j		LDEQNRDWQGLLENLHVELTLDEEDSEGPEK
						EGEGQTYFTSSKTLCGIVPQNTDSLGETGIHIC
187	1537	A	2158	227	442	AHDKSP
167	1337	1	2130	221	442	FNCFRVASDSFLENSSLLIMILPLRNATQEFIIR PGAVAYTCNPSTLGGWGGWITRSGVRDOPG
						OHGGTPS
188	1538	A	2167	3	486	AHLGGAWLTQRSLGSWAAPGPARAAKEVVA
		}		,	1.00	CIPQNQKMNIWRMKTSKHLQLLSFVLGAVSP
	İ	,				AVVVPYMMVLQENGYGVEEGIPTLLMAASS
	1					MDDILAITGFNTCLSIVFSSGCARSSGSRNSKS
	1	İ				LRTPLGTICEGCDDSSIFSHLDHSSKWSSTYG
						HSGA
189	1539	A	2168	2	412	EFLSSNQITQLPNTTFRPMPNLRSVDLSYNKL
	1	ĺ				QALAPDLFHGLRKLTTLHMRANAIQFVPVRIF
		1				QDCRSLKFLDIGYNQLKSLARNSFAGLFKLTE
	}]]			LHLEHNDLVKVNFAHFPRLISLHSLCLRRNKV
190	1540	A -	2179	64	399	AIVVSSLDW MRLNQNTLLLESFGXXRPYTSEHAPTYHOW
170	1340	^	21/3	04	377	MKADELLRWTTSEPLTLEHEYAMQRTWLED
	l					AYECTFIVLDAEKRHAQPGATEESCMVGDVN
	i	ļ	•			LFLTDLEDLTLGEIEVLIAEP
191	1541	Α	2190	1	469	CLDRAAGIRHERNVIYINETHTRHRGWLARR
						LSYVLFIQERDVHKGMFATNVTENVLNSSRV
						QEAIAEVAAELNPDGSAQQQSKAVNKVKKK
	1] [1	AKRILQEMVATVSPAMIRLTGWVLLKLFNSF
	1					FWNIQIHKGQLEMVKAATETNLPLLFLPVHR
165	1	ļ	لبييا			SH
192	1542	A	2197	26	157	PSKXGGIRLLLTGTQLYGRFGSAIAPLGDLDR
102	1647	<u> </u>				DGYNGEGREEPY
193	1543	A	2236	2	383	EYFPNSIWRSLFSTMDLGDIGFYTYRILQALS
	1			İ		YTHSKGIMHRDVKPLNILCNSPRNKVILADW
			j			GLAEFYHPMRKYSVHVATRYYKSPEILLDYE
						YYDYSLDIWAVGVILLELLTLKLHVFEGGDN
194	1544	Ā	2241	105	409	EQ RKGVGKMPTSEGRPGQERSDWVTSYKVMGS
	1 ****	^	****	103	707	NDSHTWVTVKNGSGDMIFEGNSEKEIPVLNE
	1					I TANDITE ALLA VILLO CONTRINE CILO EVEILA FUE
			i i			
						LPVPMGARYIRINPQSWFDNGSICMRMEILGC
195	1545	A	2245	1	672	

SEQ ID	T SEO ID	Met	SEQ	Predicted	Predicted end	T 4 : : 1
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	[USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1		ļ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
1.				peptide	•	/=possible nucleotide deletion. \=possible
L		<u> </u>		sequence		nucleotide insertion
						KLCGASSGIIDLLPSPSAATNWTAGLLVDSSE
	ĺ					MIFKFDGRQGAKIPDGIVPKNLTDQFTITMW
İ			1			MKHGPSPGVRAEKETILCYSDKTEMNRHHY
İ	1	1	ì	i	1	ALYVHNCRLVFLLRKDFDQADTFRPAEFHW
1	1				1	KLDQQALAKVDGQPGKSITRQLQEMPVTIQG
196	1546	-	2256	ļ. <u>.</u>		ISLKPS
196	1340	A	2236	1	396	FRGTPVSGLTNRDTLAVIRHFREPIRLKTVKP
	ł					GKVINKDLRHYLSLQFQKGSIDHKLQQVIRD
]						NLYLRTIPCTTRAPRDGEVPGVDYNFISVEQF KALEESGALLESGTYDGNFYGTPKPPAEPSPF
İ	ľ	ļ	ł	ĺ	•	OPDPV
197	1547	A	2259	43	594	QLAIEIGVRALLFGVFVFTEFLDPFQRVIQPEEI
10,	1347	"	2239	13	JJ4	WLYKNPLGQSDNIPTRLMFAISFLTPLAVICV
						VKIIRRTDKTEIKEAFLAVSLALALNGVCTNTI
	1	1	ł	ł	1	KLIVGRPRPDFFYRCFPDGVMNSEMHCTGDP
				ļ	ĺ	DLVSEGRKSFPSIHSSFAFSGLGFTTFYLAGKL
					1	HCFTESGRGKSWRLCAAILPL
198	1548	Α	2275	3	404	TCTTVVVIPRMLVDFLSESKTISLPECATOMFF
	ļ			İ	-	FLGFASNNCFIMAAMSYDRYTAIHNPLOYHT
		Ì	1			LMTRKICLQMMMASWMVGFLFSLCIIVTVFN
ĺ	l	ĺ				LSI.CDLNTIQHYFCDISPVVSLACNYTFYHEM
		<u> </u>	ļ			AIFVLSA
199	1549	Α	2315	1	375	LTQMFFIHALSAIESTILLAMAFDRYVAICHPL
		1				RHAAVLNNTVTAQIGIVAVVRGSLFFFPLPLLI
				•		KRLAFCHSNVLSHSYCVHQDVMKLAYADTL
200	1550	A	2334	2 .	409	PNVVYGLTAILLVMGXDRMFISLSYFLII
200	1550	A	2554	2 .	409	PRVRPQQRKMSFFFKTELGEKLVTKFLFETDF
		ļ				SDDPMLPSPDQLKKKAPFTNKKLKAHQTPVD
	1	1	1	•		ILKQKAHQLASMQVQAYNGGNANPRPANNE EEEDEEDEYDYDYESLSDDNILEDRPENKSCH
			1 1			DOLOFEYKEEM
201	1551	A	2350	3	512	ISWEAQIAEIIQWVSDEKDARGYLQALASKM
1	1	1		_		TEELEALRSSSLGSRTLDPLWKVRRSQKLDM
			1			SARLELQSALEAEIRAKQLVQEELRKVKDAN
	i ,					LTLESKLKDSEAKNRELLEEMEILKKKMEEK
1		l	i i			FRADTGKLMLCDSALFEYKYFSNECFYFLFD
						LIVTLEAPTEFQIQY
202	1552	A	2351	1	1003	PSSYSSDELSPGEPLTSPPWAPLGAPERPEHLL
				l		NRVLERLAGGATRDSAASDILLDDIVLTHSLF
					}	LPTEKFLQELHQYFVRAGGMEGPEGLGRKQA.
					1	CLAMLLHFLDTYQGLLQEEEGAGHIKDLYL
	. '				i	LIMKDESLYQGLREDTLRLHQLVETVELKIPE
				ļ		ENQPPSKQVKPLFRHFRRIDSCLQTRVAFRGS
	-			-		DEIFCRVYMPDHSYVTIRSRLSASVQDILGSV TEKLQYSEEPAGREDSLILVAVSSSGEKVLLO
						PTEDCVFTALGINSHLFACTRDSYEALVPLPE
						EIQVSPGDTEIHRVEPEDVANHLTAFHWELFR
						CVHELEFVDYVFHGE
203	1553	Α	2361	2	403	NNLNCAEPLFEQNNSLNVNFNTQKKTVWLIH
ĺ					ŀ	GYRPVGSIPLWLQNFVRILLNEEDMNVIVVD
	l				1	WSRGATTFIYNRAVKNTRKVAVSLSVHIKNL
	_ [İ	LKHGASLDNFHFIGGSLGAHISGFVGKIFHGO
			{			LGRITGLDP
204	1554	A	2390	280	476	SPSLLPQCLMSLSDLSLSPAPPSHLSPRCPSPQ
				J]	AGSRLGAMRRCAREMDATPMPPAPSCPSERV
205	1555		0.400			T
205	1555	A	2400	543	745	AAVALRDISWQQPYPMDFYAGSSLGPWTVN
				ļ	Ì	HGQDRRPHAPGRPARGKVQEGSARPPSAVAC
					1	EDCSCR

CEO ID	SEQ ID	Mat	LCEA	l Deading	1 8 20 1	
SEQ ID NO: of	NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	noa	in NO.	nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide	, , ,	1			location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq-	ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	1	}	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i	1	l	1	peptide		/=possible nucleotide deletion, \=possible
		<u> </u>		sequence		nucleotide insertion
206	1556	A	2406	122	485	DLSPDSREDHPQGHRRLLPKRPVRGSLMPGH
		1	1] .	THHPCPVSSTTNDTPDQIWVSVGSLRMGTGG
1	1	1				MGANASTSPRCWDLSSGNKKWIIQVPILASIV
	1			<u> </u>	}	ESRGGLLATGVGGMCACVPRNQPLTGT
207	1557	Α	2409	289	418	LWTLYRHKQQVQHNHSNRLSCRPSQEDRAT
L	1	ł	1		1	HTIMVLDKENTLS
208	1558	A	2413	64	492	VQGTGXXFIAFTEAMTHFPASPVWAGMFFL
!	l		ł		}	MLINLGLGSMIGTMAGITTPIIDTFKVPKEMFT
	!	1				GGCCVFAFLVGLLFVQRSGNYFVTMFDDYSA
1	i					TLPLTLIVILENIAVAWIYGTKKFMQELTEML
i	ľ	l			ł	GFRPYRFYFYMWKFVSP
209	1559	A	2417	3	877	EKERLLDEWFTLDEVPKGKLHLRLEWLTLMP
1				-	1	NASNLDKVLTDIKADKDQANDGLSSALLILY
ł	}	1			}	LDSARNLPIRYKTNEPVWEENFTFFIHNPKRO
		1				DLEVEVRDEQHQCPLGNLKVPLSQLLTSEDM
ł	1	l				
	1				1	TVSQRFQLGNSGPNSTIKMKIALRVLHLEKRE
1	Į.	i			f	RPPDHQHSAQVKRPSVSKEGRKTSIKSHMSG
1	ŀ	1				SPGPGGSNTAPSTPVIGGSDKPGMEEKAQPPE
i	1					AGPQGLHDLGRSSSSLLASPGHISVKEPTPSIA
1]	1	1 1		Į	SDISLPIATQELRQRLRQLENGTTLGQSPLGQI
210	1560	<u> </u>	0422	25	150	QLTIP
210	1300	A	2422	35	456	REFAASDLEPFTPTDQPISPEAITQPSCIKRQRA
ŀ	ļ)] ']			AGNPGSLAATIDHKPCSAPLEPKIQASRNQRW
1	1		1			GAVRAAESLTDIAEPASPQVHETPIDASQTQK
i	l		1			VEPASKSRFTPELQAKVSHSRERALSTMDATP
	<u> </u>	<u> </u>				HHAQPQRGEG
211	1561	Α	2431	1 .	764	RRYSQKLIQHTACQLLRTYPAATRIDSSNPNP
]	J	j				LMFWLHGIQLVALNYQTDDLPLHLNAAMFE
						ANGGCGYVLKPPVLWDKNCPMYQKFSPLER
1	ŀ					DLDSMDPAVYSLTIVSGQNVCPSNSMGSPCIE
		l				VDVLGMPLDSCHFRTKPHRNTLNPMWNEQF
i i		1				LFHVHFEDLVFLRFAVVENNSSAVTAQRIIPL
!						KALKRGYRHLQLRNLHNEVLEISSLFINSRRM
1						EENSSGNTMSASSMFNTEERKCLQTHRVTVH
						GVPG
212	1562	Α	2436	1	411	GIRGTTGHLGCPINDDPSLTLTVSWVMEDKPI
					İ	YIGNGTKKEDDSLTIFAVAKRDHVSDTCGAC
						TDLDHNLDKGYLTVLGEQATPTNRLGALPKG
						RANRTRDLELTYLAERIVRLTWIPGDANNRPI
L					j	TDYDCQIEEHQ
213	1563	Α	2445	1	1294	MSSIGCLWVSRSSQIDGLTAEKSGPEKPHGT
			}	ļ		WLMPELHPKEQILELLVLEQFLSILPEELQIWV
			ı İ	ļ		QQHNPESGEESVTLLEDLEREFDDPGQQVPAS
		ĺ				PQGPAVPWKDLTCLRASQESTDIHLOPLKTO
			-			LKSWKPCLSPKSDCENSETATKEGISEEKSOG
						LPQEPSFRGISEHESNLVWKQGSATGEKLRSP
					•	SQGGSFSQVIFTNKSLGKRDLYDEAERCLILT
				İ		TDSIMCQKVPPEERPYRCDVCGHSFKOHSSLT
ľ				i	ł	
					1	QHQRIHTGEKPYKCNQCGKAFSLRSYLIIHQR
				į		IHSGEKAYECSECGKAFNQSSALIRHRKIHTG
			. []		EKACKCNECGKAFSQSSYLIHQRIHTGEKPY
		1		l	ļ	ECNECGKTFSQSSKLIRHQRIHTGERPYECNE
ļ		1			1	CGKAFRQSSELITHQRIHSGEKPYECSECGKA
214	1566					FSLSSNLIRHQRIHSG
214	1564	A	2461	1	615	GIPGSTISSSRNIFLEDDLAWQSLIHPDSSNTPL
		1		J		STRLVSVQEDAGKSPARNRSASITNLSLDRSG
ļ	1		1		1	SPMVPSYETSVSPQANRTYVRTETTEDERKIL
l	ļ	ł	i	}	j	LDSVQLKDLWKKICHHSSGMEFQDHRYWLR
1						THPNCIVGKELVNWLIRNGIIIATRAQAIAIGQ

CCC 15	10000	137:	LOCO	I Design	T-6 11	T
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	пои	in NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	"""]	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1	•	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ļ			ļ	peptide	1 304	/=possible nucleotide deletion, \=possible
			1	sequence	İ	nucleotide insertion
	<u> </u>					AMVDGRWLDCVSHHDQLFRDEYALYRPLQV
		ļ	1	i	ļ	LFSVYCQLECSKLIL
215	1565	Α	2464	3	2932	GPGVRSSQDGMADVFVHLRTAWPRCSFISGQ
Į				1		HGPGRHGRRVCSSODSMADVFVHLRTAWPT
ł	i	ł		1	1	CSLISGQHGPGESVSYEDDDIPAPASLLHVNA
l		1			J	AAPALTNPTAPVLCTAPNNTAQKEKVPSGMR
		İ		ł		QRPAGVRISSRTPDLTCAVSTHSTVPGVRISSC
ľ	1		1		ł	TPDLTCAVSIHSTVPSVCISSCTPDLTCAVSTH
				}		STVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
	j					TPDLTCAVSIHATVPGVRISSCTPDLTCAVSIH
	ļ					ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
	Ì					TPDLTCAVSIHSTVPGVRISSCTPDLTCAVSIH
	ļ]	1			ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
	1	l	ŀ			TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH
	ŀ		1			ATVPGVRISSCTPDLTCAVSIHATVPGVRISSC
	1		1	}	ļ	TPDLTCAVSIHATVPGVRISSRTPDLTCAVSIH
İ						ATVPGVRISSCTPDLTCAVSTHSTVPGVRISSR
		i				TPDLTCAVSIHATVPGVRISSCTPDLTCAVSTH
l			1			STVPGVRISSRTPDLTCAVSIHATVPGVHISSC
İ		1	1			TPDLTCAVSTHSTVPGVRISSRTPDLTCAVSIH
!]			STVPGVCISSRTPDLTCAVSIHSTVPSVHISSCT
ì	ł		1			PDLTCAVSIHSTVPGVRISSRTPDLTCAVSTHS
	ŀ	ĺ	1			TVPGVHISSCTTDLTCAVSHATVPGVHISSCT
	ļ					PDLTCAVSTHTTVPGVRISSRTPDLTCAVSIHS
ł	}	1	ľ			TVPGVRISSCTPDLTCAVSTHSTVPGVRISSRT
ļ						PDLTCAVSTHLTVPGVRISSRTPDLTCAVSIHA TVPGVHISSCTPDLTCAVSIHATVPGVRISSRT
				•		PDLTCAVSIHATVPGVHISSCTPDLTCAVSTHS
ĺ		i	1			TVPGVRISSRTPDLTCAVSIHSTVPGVHISSCT
1			}	,		PDLTCAVSTHSTVPGVHISSCTPDLTCAVSTH
	1					STVPGVHISSRTPDLTCAVSIHATVPSVHISSC
ĺ	Ĭ					TPDLTCAVSIHSTVPGLLTSVSQTSTG
216	1566	Α	2477	1	414	FRTKSYRKGSYRCIVSEWIAEQGNWQEIQEK
ļ	j					AVEVATVVIQPTVLRAAVPKNVSVAEGKELD
						LTCNITTDRADDVRPEVTWSFSRMPDSTLPGS
						RVLARLDRDFLVHSSPHVALSHVDARSYHLL
ł	1	}	1	ļ ,		VRDVSKENSGYYY
217	1567	A	2480	2	460	CRTLCEGPQRFEEYEYLGYKAGLYEAIADHY
						MQVLVCQHECVRELATRPGRLSPIENFLPLHY
	1		1 1			DYLQFAYYRVGEYVKALECAKAYLLCHPDD
			[]	ļ		EDVLDNVDYYESLLDDSIDPASIEAREDLTMF
<u></u>				i		VKRHKLESELIKSAAEGLGXSYTEPNYW
218	1568	A	2483	140	383	AFSSPHPSPAPQFPECGFYGLYDKILLFKHDPT
		!				SANLLQLVRSSGDIQEGDLVEVVLSASATFED
		-		-		I.QIRPHALTVHSYRAP
219	1569	Α	2489	3	428	SSRLVLLAGAAALASGSQGDREPVYRDCVLQ
				Į.		CEEQNCSGGALNHFRSRQPIYMSLAGWTCRD
. [[DCKYECMWVTVGLYLQEGHKVPQFHGKWP
]		1 1	}	ļ	FSRFLFFQEPASAVASFLNGLASLVMLCRYRT
			1	i		FVPASSPMYHTCVAFAWVS
220	1570	A	2498	1	1297	MDGEAVRFCTDNQCVSLHPQEVDSVAMAPA
			[ľ	ł	APKIPRLVQATPAFMAVTLVFSLVTLFVVDH
	[.		HHFGREAEMRELIQTFKGHMENSSAWVVEIQ
	[1		MLKCRVDNVNSQLQVLGDHLGNTNADIQMV
}	1		1 1	ľ		KGVLKDATTLSLQTQMLRSSLEGTNAEIQRL
					T. C. C. C. C. C. C. C. C. C. C. C. C. C.	
	1			1		KEDLEKADALTFQTLNFLKSSLENTSIELHVL
			ļ			SRGLENANSEIQMLNASLETANTQAQLANSS
	:					

SEQ ID	SEQ ID	Mct	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide	====	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ĺ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1 .	l	١	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
	ļ			peptide	Sequence	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
ļ				Sequence		FDNTSAEIQFLRGHLERAGDEIHVLKRDLKM
						VTAQTQKANGRLDQTDTQIQVFKSEMENVN
1					İ	TLNAQIQVLNGHMKNASREIQTLKQGMKNA
	i	l	l	Ì	ŀ	SALTSQTQMLDSNLQKASAEIQTLRQDLENT
		ŀ				KALTMEIQQEQSRLKTLHVVITSQEQLQRTQ
221	1571	A	2501	3	500	RVRLNNDGLSPLMMAAKTGKIGIFQHIIRREV
221	13/1	A	2301] 3	1 200	TOEDTRIII ODVEVDAVANCOLANDI OOLD
	l			ì		TDEDTRILLSRKFKDWAYGPVYSSLYDLSSLD
			Ĭ	ĺ		TCGEEASVLEILVYNSKIENRHEMLAVEPINE
ŀ	l	i			1	LLRDKWRKFGAVSFYINVVSYLCAMVIFTLT
1	l			ļ		AYYQPLEGTPPYPYRTTVDYLRLAGEVITLFT
			L			GVLFFFTN
222	1572	A	2508	3	395	DAHCQRKLAMQEFMEINERLTELHTQKQKL
1	1					ARHVRDKEEEVDLVMQKVESLRQELRRTER
	l	}	1			AKKELEVHTEALAAEASKDRKLREQSEHYSK
1		1	1			QLENELEGLKQKQISYSPGVCSIEHQQEITKL
	L	<u>L</u>	L		L	KTDLEKKS
223	1573	A	2544	2	412	NDPAIISNFSAAVVHTIVNETLESMTSLEVTK
	ļ		1			MVDERTDYLTKSLKEKTPPFSHCDQAVLQCS
1	}		ł		İ	EASSNKDMFADRLSKSIIKHSIDKSKSVIPNID
Ì		l				KNAVYKESLPVSGEESQLTPEKSPKFPDSQNQ
	ļ	l				LTHCSLSAA
224	1574	A	2552	401	1	GASLCFISTAFTVLTFLIDSCRFSYPERPIIFLSM
		**			_	CYNIYSIAYIVRLTVGRERISCDFEEAAEPVLI
		1				QEGLKNTGCAIIFLLMYFFGMASSIWWVILTL
			1			TWFLAAGLKWGHEAIEMHSSYFHIAAWAIPA
1						VK .
225	1575	A	2563	724	1	MSARKERREKGEEEGEGEKDGDEDEKEEEKE
	13/3	1	2505	124	'	GLGEEEKEAGKKKKKQEEKEKGAVYSR
l	İ	ĺ	1			VARICKNDMGGSQRVLEKHWTSFLKARLNC
		İ	ļ	•		SVPGDSFFYFDVLQSITDIIQINGIPTVVGVFTT
		l				QLNSIPGSAVCAFSMDDIEKVFKGRFKEQKTP
i		l				DSVWTAVPEDKVPKPRPGCCAKHGLAEAYK
ŀ						1
[(1	1	·		TSIDFPDETLSFIKSHPLMDSAVPPIADEPWFT
226	1576	<u> </u>	2571	440		KTRVRYRLTAISVDHSAGPYH
220	13/6	A	2571	449	3	EGVLFVYGNYVGDVMNFEMAAEMAQEVAIP
}	ł	1			,	TRTVLTTDDISSSPIEDRDGRRGVAGNFFIFKV
'	1	1		,		AGAACDRGMSLEACEAVTRKANRRTYTMG
			1			VALEPCSLPQTRRHNFEIGAEEMEIGMGIHGE
		 			1.00	RGVIREKMMPADAIVDHIMDRIFS
227	1577	A	2575	3	1197	VLSDLCLFYYRDEKEEGILGSILLPSFQIALLTS
]				EDHINRKYAFKAAHPNMRTYYFCTDTGKEM
	1					ELWMKAMLDAALVQTEPVKRVDKITSENAP
	Į.	1	,			TKETNNIPNHRVLIKPEIQNNQKNKEMSKIEE
}	ļ].			KKALEAEKYGFQKDGQDRPLTKINSVKLNSL
		i				PSEYESGSACPAQTVHYRPINLSSSENKIVNVS
		1				LADLRGGNRPNTGPLYTEADRVIQRTNSMQQ
	1	ļ)	ļ		LEQWIKIQKGRGHEEETRGVISYQTLPRNMPS
		ì				HRAQIMARYPEGYRTLPRNSKTRPESICSVTP
	!	l	†			STHDKTLGPGAEEKRRSMRDDTMWQLYEW
	1	[i			QQRQFYNKQSTLPRHSTLSSPKTMVNISDQT
	1	1				MHSIPTSPSHGSIAAYQGYSPQRTYRSEVSSPI
		l				ORGDVTIDRRHRAHHPKVK
228	1578	A	2583	3	330	LPFLGLGSVLPOGMVMASPEMNPTICSVFEA
	15.0	l "`	2505	ا	.	HIVLLFHATTFRRGFQVTVLVGNVRQTAVVE
}			l '	·		KIHAKVRGTWPFISPEVRKEGGLPQTGRELLD
	[1	[
229	1570	 	2590	1	110	PTMGIKPHLWWVAA
229	1579	A	2589	•	448	DDKNAQGIKRHVKPTSGNAFTICKYPCGKSR
1	1	ĺ	[ECVAPNICKCKPGYIGSNCQTALCDPDCKNH
						GKCIKPNICQCLPGHGGATCDEEHCNPPCQH

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	İ			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
]	ł		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ľ		ì	}	peptide		/=possible nucleotide deletion, \=possible
<u> </u>		 		sequence		nucleotide insertion
						EMANKELKQLRASYTESCIQEHYLPQVIDGTL Y
241	1591	A	2640	392	3	IRLTILRCVFMRLATICVLVFTLGSKITSCDDD
		ļ				TCDLCGYNQKLYPCWETQVGQEMYKLMIFD
		ĺ	[FIIILAVTLFVDFPRKLLVTYCSSCKLIQCWGQ
						QEFAIPDNVLGIVYGQTICWIGAFFSPLLPAM Y
242	1592	Α	2642	405	1	YFKNTTLLLVGVICVAAAVEKWNLHKRIALR
242	1392	Λ.	2042	403	1	MVLMAGAKPGMLLLCFMCCTTLLSMWLSNT
					İ	STTAMVMPIVEAVLQELVSAEDEQLVAGNSN
			}			TEEAEPISLDVKNSQPSVELIFVNEDILDFLMK
						SPLMISQACI
243	1593	A	2646	412	2	CLAMIKGIQSSGKIIYFSSLFPYVVLICFLIRAF
						LLNGSIDGIRHMFTPKLEIMLEPKVWREAATQ
<u> </u>						VFFALGLGFGGVIAFSSYNKRDNNCHFDAVL
]			VSFINFFTSVLATLVVFAVLGFKANVINEKCIT
	-					QNSETV
244	1594	Α	2650	1	1271	MTTTLIGLLKTARLLRLVRVARKLDRYSEYG
1						AAVLMLLMCIFALIAHWLACIWYAIGNVERP
						YLTDKIGWLDSLGQQIGKRYNDSDSSSGPSIK
						DKYVTALYFTFSSLTSVGFGNVSPNTNSEKIF SICVMLIGSLMYASIFGNVSAIIQRLYSGTARY
					·	HMQMLRVKEFIRFHQIPNPLRQRLEEYFQHA
						WTYTNGIDMNMVTNGTCSSCTSDDGHFiLVS
[NHHQGGLIYSWNDAASMQRPFNHIKSSLLGS
						TSDSNLNKYSTINKIPQLTLNFSEVKTEKKNSS
					1	PPSSDKTIIAPKVKDRTHNVTEKVTQVLSLGA
						DVLPEYKLQAPRINKFTILHYSPFKAVWDWLI
]				, "		LLLVIYTAIFTPYSAAFLLNDREEQKRRECGY
i !						SCSPLNVVDLIVDIMFIIDILINFRTTYVNQNEE
245	1595		3666	205		VVSDPASV
245	1393	Α	2656	385	2	NLTWWPLFRDVSFYIVDLIMLIIFFLDNVIMW
				·		WESLLLLTAYFCYVVFMKFNVQVEKWVKQ MINRNKVVKVTAPEAQAKPSAARDKDEPTLP
1				ĺ		AKPRLORGGSSASLHNSLMRNSIFONKIHTLD
	'					PHV
246	1596	Α	2660	200	506	VLVLOMNYYOMLIIYYVLFFKVNEFLAFEGPI
						LLDMRIKHLIKTNQLSQATALAKLCSDHPEIG
						IKGSFKQTYLVCLCTSSPNGKLIEEVSMFSFIS
						NYFLS
247	1597	Α	2678	3	267	DAWVKNDIIFNQTERKQKISENLKHLASVRV
}		ļ		j		VQKNLVFVVGLSQRLADPEVSPLVFFVILIFF
248	1598	A	2687	1	404	VSLSYLEIIFDPAQLCDSSEHIIS
240	1370	Δ.	200/		.404	DFTTLAAMMRTLFSLFGDVRSDVHRFSVTLF GAAIKSVKNPDKKSIENQVLDSLVPLLLYSQD
	ŀ	l	J			ENDAVAEESRQVLTICAQFLKWKLPREVYSK
'	ŀ		!			DPWHIKPTEAGTICRFFEKKCKGKINILEQTL
	ļ	1]		ļ	MYSKNPKL
249	1599	A	2692	ī	440	FRRRRRRERDCAAQGARRHCRHLAECKLV
						SFPIGIYKVLRNVSGQIHLITLANNELKSLTSK
	1		1			FMTTFSQLRELHLEGNFLHRLPSEVSALQHLK
	{		[[AIDLSRNQFQDFPEQLTALPALETINLEENEIV
						DVPVEKLAAMPALRSINL
250	1600	Α	2693	459	21	LLPGSLGVPILHSQPWDPSPQCPHRAPSTPRRL
	ļ	1	ł			PPLGALSQALTFLSRAAKNHSQDPGKGTKPFP
		- 1	J	1		AAPAAPPPRSSLPAPLPMGLKDKGPQPAPPTIF
1		- 1	- 1		ĺ	NSPWHPATLPGALGPQLSQAAPSPIPPPCLMG ISSCPDLKLTKSSTP
251	1601	A	2694	2	404	FVFDLKLRVPGFAALLIHGASSVPGPETVRLR
						1 TI DEREKT GEARDERINGAGS TOTE I VICE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:			
	1	поа	1	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	scq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ſ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		l	l	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
Į.	ŀ	Ī		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	i	ſ	[peptide	1	/=possible nucleotide deletion, \=possible
1]		ŀ	sequence	1	nucleotide insertion
	-	 	 	Sequence		
1	ł	ĺ]			QKRKKKAPDHSSGRKEELVTTHTVDKLETKK
1	}]		J	PVGRVLCGLSGELLHSLLLPRRKTEKRALGSH
		1				RKAGFPEHPVAPEPLSNSCQISKEGREQVLSEI
			l		j	GAGDCL
252	1602	Α	2697	421	1	PQKSHSGAYQCFATRKAQTAQDFAIIALEDG
1	ļ		1]	TPRIVSSFSEKVVNPGEQFSLMCAAKGAPPPT
ì	1	ł	1] .	VTWALDDEPIVRDGSHRTNQYTMSDGTTISH
		i				MNVTGPQIRDGGVYRCTARNLVGSAEYQARI
	ļ		1			
253	1603		2600		401	NVRGPPSIRAMRNIT
253	1003	Α	2698	65	401	ACCQWRRTLIPAKSTTVSCTISTPHHPFRGSYS
1	ľ	l	i			FDDHITDSEALSRSSHVFTSHPRMLKRQPAIEL .
	i			•		PLGGEYSSDVPRPLSTQLSSSLLGYFSTLMTG
Į į		1	j			AAFTNNIASSTIIL
254	1604	Α	2699	438	301	GQIHSQDDPPFIDQLGFGVAPGFQTFVACQEQ
		l	•			RVRGPWEAGPGVGY
255	1605	A	2700	1	842	I
	1003	۱^	2,00	•	042	LQNREDSSEGIRKKLVEAEELEEKHREAQVS
ļ		l				AQHLEVHLKQKEQHYEEKIKVLDNQIKKDLA
1		1				DKETLENMMQRHEEEAHEKGKILSEQKAMIN
i .		1				AMDSKIRSLEQRIVELSEANKLAANSSLFTQR
'			ŀ			NMKAQEEMISELRQQKFYLETQAGKLEAQN
1 .		1				RKLEEQLEKISHQDHSDKNRLLELETRLREVS
						LEHEEQKLELKRQLTELQLSLQERESQLTALO
		ľ	1			AARAALESQLRQAKTELEETTAEAEEEIQALT
[]			1			
256	1606		0201		10.5	VGLGSNIFRLLKASARMSVELALSILAHP
236	1000	A	2701	2	405	FVGGPGADPPVAVMWDPRAARMDLTAYAE
						LLKESGNQVLKNGNFSLAIRKYDEAIQILLQL
]			YQWGVPPRDLAVLLCNKSNAFFSLGKWNEA
l 1						FVAAKECLQWDPTYVKGYYRAGYSLLRLHQ
1.						PYEAARMFFEGLR
257	1607	Α	2702	2	399	FVESASSRPPGCFSGDGRFWLVSEGSRRGWD
i i				-	• • • • • • • • • • • • • • • • • • • •	FNPSFSFLDPRYSVGGDENIGTVTTLANILREF
ł • ł			! !			ADDITACTOR CONTROLLAR AND ANA CORA-
				į		NPSLKGFSVGTGKETSPNAFLNQAVAGGRAE
, ,						DLPVQARRLVDLMKNDTRIHFQEDWKIITLFI
350-	1600		2000			GGNDL
258	1608	Α	2709	1	1097	SVGARQGEARDRIRRFFPKGDLEVLQAQVERI
1	i				}	MTRKELLTVYSSEDGSEEFETIVLKALVKACG
1	•			ĺ	1	SSEASAYLDELRLAVAWNRVDIAQSELFRGDI
						QWRSFHLEASLMDALLNDRPEFVRLLISHGLS
						LGHFLTPMRLAQLYSAAPSNSLIRNLLDQASH
	·			ł	ł	SAGTKAPALKGGAAELRPPDVGHVLRMLLG
				ļ.		
]]				ļ		KMCAPRYPSGGAWDPHPGQGFGESMYLLSD
					ſ	KATSPLSLDAGLGQAPWSDLLLWALLLNRA
, ,					j	QMAMYFWEMGSNAVSSALGACLLLRVMAR
[[[į		LEPDAEEAARRKDLAFKFEGMGVDLFGECYR
	{		1		-]	SSEVRAARLLLRRCPLWGDATCLQLAMQAD
{	1			Į.	ļ	ARAFFAQDGVQSLPTQKWWGDMARR
259	1609	A	2721	1	403	VYLGAGPGLFFSNEGAKEGEKANIPKLMLPR
]					GGFSOREMVTGERSPSPEEEEEEEEEGFGERA
	1		[_ [
	- 1			-		SCRRGLFRVRLTRVGLAAPSKASRGQEGDAA
	1					PKSPVREKSPKFRFPRVSLSPKARSGSGDQEE
	4.55					GGLRVRLP
260	1610	A	2728	1	477	LLGGDLRYHLQQNVHFTEGTVKLYICELALA
		1		ļ		LEYLQRYHUHRDIKPDNILLDEHGHVHITDFN
		l		!	İ	IATVVKGAERASSMAGTKPYMAPEVFOVYM
	•	[ĺ	Ī	. 1	DRGPGYSYPVDWWSLGITAYELLRGWRPYEI
		1	i l	1	ł	
		1	· }	1	J	HSVTPIDEILNMFKVERVHYSSTWCKGMVAL
261	1611		2720		547	LRK
401	1011	A	2730	3	547	LTITDFILVLYRYYRSPLVQIYEIEQHKIETWR
L						EIYLQGCFKPLVSISPNDSLFEAVYTLIKNRIH

SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	[USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
""	ı	[1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	-	į		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ł	1	l	1	peptide	sequence	/=possible nucleotide deletion, \=possible
		Į.		sequence		nucleotide insertion
 	 	 	 	Sequence	}	DI DILI DDILIONI
ì	1	ł	1			RLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLP
		1			i	RPSFLYRTIQDLGIGTFRDLAVVLETAPILTAL
1	į	ĺ	ł		1	DIFVDRRVSALAVVNECGTHPQDERLGLGW
262	1612		3722			GLGEPGSEERLFPAAITSR
202	1012	A	2733	3	431	GPEFPGSAKLVFLDLSYNNLTQLGAGAFRSA
Ì					1	GRLVKLSLANNNLVGVHEDAFETLESLQVLE
			1			LNDNNLRSLSVAALAALPALRSLRLDGNPWL
[1	l	1	ĺ	Í	CDCDFAHLFSWIQENASKLPKGLDEIQCSLPM
	 	ļ				ESRRISLRACRRPASRV
263	1613	A	2736	2	343	PARISGVDPPVRKATKGGENCSFEDNKNWQF
1		i		Ì	İ	LWGLNGNFNFFKEPWGGRNNHAKGFRTTW
		ŀ				ARSSSQNNRTFQNNRNFLRLQRDSQKKGQFA
]_	RLISPLVNLPQSPGGLEFQYQAT
264	1614	A	2738	2	245	RAMLKCLREGQPPPSYNWTRLDGPLPSGVRV
	:	1	1			DGDTLGFPPLTTEHSGIYVRHDTNEFSSRDSH
L	1	ĺ	[DTVDVLDPPEDSGKQVDL
265	1615	Α	2752	2	388	AAGDAPLRSLEQANRTRFPFFSDVKGDHRLV
	ł	1		,		LAAVETTVLVLIFAVSLLGNVCALVLVARRR
	1					RRGATACLVLNLFCADLLFISAIPLVLAVRWT
	1					EAWLLGPVACHLLFYVMTLSGSVTILTLAAV
						SLER
266	1616	Α	2755	192	1	AFREVGGYWGLLCEHLYAIPSKTSEGNWTAK
[-	LQGYLPLQDAFHIFQDPLTGDLPWPELILGLP
						V
267	1617	Α	2760	434	714	ASRLEKQNSTPESDYDNTPNDMEPDGMGYM
	1		1 2700		7.11	HRTSVPGEGLPRARDLAGLGQQKQFTTHTPF
						LYFQTHKGLKDSSIRSEVTCLGISQCWRKGFF
268	1618	Ā	2762	1	405	IACTFCGQDEWSPERSTRCFRRSRFLAWGEP
					705	AVLLLLLLSLALGLVLAALGLFVHHRDSPL
	1 1					VQASGGPLACFGLVCLGLVCLSVLLFPGQPSP
					ľ	ARCLAQQPLSHLPLTGCLSTLFLQAAEIFVESE
•	1 1					LPLSWAE
269	1619	A	2772	3	243	
207	1.0.7	^	2112	,	243	TRPAEKIQYLVLFFVMSHPSQAYDKLSLSDHL
))					LIAVLNLLRREVSEHGRHLQQYFNLFVMYAN
270	1620	A	2789	1	496	LSKNLSFSEFCFDVSY
210	1020	^	2109	•	486	ELQSQQACTHTKETEQLRSQLQTLKQQHQQA
	[(VEQIAKAEETHSSLSQELQARLQTVTREKEEL
						LQLSIERGKVLQNKQAEICQLEEKLEIANEDR
		Į	l		ļ	KHALERFEQEAVAVDSNLRVRELQRKVDGIQ
		j		}		KAYDELRLQSEAFKKHSLDLLSKERELNGKL
271	1.00		2505			RHLSP
271	1621	A	2795	1	568	KEKRVTVQLPTESIQKNQEDKLKMVPRKQRE
					1	FSGSDRGKLPGSEEKNQGPSMIGRKEERLITE
		- 1	- 1		ļ	RKHEHLKNKSAPKVVKQKVIDAHLDSQTQN
	1 1	ł		- 1	ŀ	FQQTQIQTAESKAEHKKLPQPYNSLQEEKCLE
		- 1	ľ	ł		VKGIQEKQVFSNTKDSKQEITQNKSFFSSVKE
]	SQRDDGKGALNIVEFLRKREELHQILSTVKQP
272	1622	A	2797	8	523	KCMQGKYAGAMESEPCVCTEADFDCDYGYE
]]	. 1	1		ĺ	RHSNGQCLPAFWFNPSSLSKDCSLGQSYLNST
	[- 1	ľ	i	l	GYRKVVSNNCTDGVREOYTAKPOKCPGKAP
		- 1			l	RGLRIVTADGKLTAEQGHNVTLMVQLEEGD
	1	- 1	1	ł	ł	VQRTLIQVDFGDGIAVSYVNLSSMEDGIXHV
			ļ	i		YQNXGIXRXTVQVDNSLGS
273	1623	A	2801	72	395	HPSRSNVGPRQLTVWNTSNLSHDNRRKYIFS
				·~	1	DEEGQNQLGIRIHQDIPLPPRRRELPALRTTNG
		ł	1	!		VADOLARIODA COL CELEVOTO COLOR
	, !	1	1	1	ì	KADSLNVSRNSVMQELSELEKQIQVIRQELQL
	1					
274	1624	<u>_</u>	2805	168	220	AVSRKTELEEYH
274	1624	A	2805	168	320	AVSRKTELEEYH ILWLYFETGTWVYPVFAKLSLLGLAALFSLRE IFIARNGVVGETLTHCKRV

NO: of NO: of No: of N	SEQ ID	SEQ ID	Mct	SEQ	Predicted	D-4:-4-44	14-1-14
Deput						Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Decision Decision	1		1100				D=Asparuc Acid, E=Glutamic Acid,
Sequence							r=rnenylalanine, G=Glycine, H=Histidine,
September Sept							
minn acid residue of peptide sequence peptide s			1	1		1	
Persidue of peptide peptide Petide Peptide Peptide Peptide Peptide Peptide Peptide Peptide Peptide Peptide Peptide Peptide Peptide Peptide Peptide Peptide Peptide Peptide Peptide Pepti		ļ		1			
Poptide			ĺ	ì	1		Y=Tyrosine, X=Unknown, *=Ston codon
	1	1	1		peptide	'	/=possible nucleotide deletion. \=nossible
MGKIIPO			1	1	sequence		
MGKIIPO	275	1625	A	2812	208	321	GSLATCQLSEPLLWFILRVLDTSDALKAFHD
1627		.l				İ	
KYQQTVVADLAGDETIFGSSLIPGHVQAY OYGPKRNGBAGPG	276	1626	A	2813	41	266	AGRSLHGAGDRAWVGISPTDWSPKVVELCK
277	1	1	Į	1		i	
LFISYLHTPKHKQHEVLQAMGSILGITGEEME PLFQEEHGTARRWITGURGOSKSVPKTJE			<u> </u>	<u> </u>			QVGPVRRNGEAGPG
P.F.O. P	277	1627	A	2817	3	410	VLQERLDNFQRKCIQLASSTEGKVDKLLMRN
GLNQOPALNGSFSELFVKFLKTESLSSTLPTX	1.	Į.	1			İ	LFISYLHTPKHKQHEVLQAMGSILGITGEEME
LPPHNSPOKIK			ľ	1	ĺ	1	
1628	İ						
279	020	1600	ļ.,	I			
PVDONPRLV	2/8	1628	A	2821	238	457	GLSGPSCSCPHSPLPTIISRAQLETALKWRNYE
1629		1				ļ	VKLRLLLHLEELQMEHDIRHYDLESVPMTWD
TPPTFSSVPPPLPSILSSLHHSPLHSELNPHLQS	270	1620	 	1 2022	740	<u> </u>	
CRLPSRPSVSELEPPQSGPASSVPLAPTPLPDS	219	1629	A	2822	342] 1	
PSWYWSYHWGYKOKRLALCVFSFEEGGRRK		Ì	1				TPPTFSSVPPPLPSLSSILHHSPLHSELNPHLQS
280				1			CRLPSKPSVSRELPPQSGPASSVPLAPIPLPDS
281	280	1630	_	2825	207	77	
VAFQCDGQRREFTC	200	1030	1 ^	2023	307	''	
281			İ				
282	281	1631	A	2827	RI	381	
NTTNMDEVPRPQALSGSSVVWVSGCVASRS VILSLTSG	-0.	1031	ļ "·	202.		201	
VILSLTSG							NTTNMDEVPRPOAL SGSSVVVVVSGCVASPS
1632			ļ				VILSUTSG
TSSGKYNELGYPFGYLKASTTLTCVNLFVMP	282	1632	A	2830	471	160	KLPXDKYFLEPSPLTOYILERKSPHTCWOVEV
283		į					TSSGKYNELGYPFGYLKASTTI TCVNI FYMP
YLKTLPPYYL			ł				YNYPVLLPLLDDLFKVHKLKPNLKWROAFDS
1633	Į	ļ	j]			
MKHYLHSQACSVFNYHLSPRTFPRYPGLMVP PLQCQMIPEESTQFSIKLQPPPVGRKNRERVE SSESSAP	283	1633	Α	2835	462	148	
284	i I		ł	1			MKHYLHSQACSVFNYHLSPRTFPRYPGLMVP
SSEESAP							
DYGLVVRGCLDLRYLAMRQRNNLLCNGLSL KSLAETVLNFPLDKSLLLRCSNWDAETLTED QVIYAARDAQISVALFLHLLGYPFSRNSPGEK KSLAETVLNFPLDKSLLLRCSNWDAETLTED QVIYAARDAQISVALFLHLLGYPFSRNSPGEK KR KR KR KR KR KR KR				<u> </u>			SSEESAP
285 1635 A 2843 20 271 PIRPYYSYSGLDRDCSWLPLAKAWLPDVMIL VCDRVSEDGINRQQAQEWCIKHGFELVELSP EELPEEDGKCLCVRKYGTYI 286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFTNSELHRAM NLHVGNLRLLSGPLDQVRAALPTPALSPKDK AVLQNLKRILAKVQEMRDQRVSLEQQLRELI QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLFYNLTRDETYJWEK ITDFEGQPPTPRDKLSCWYYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQQQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	284	1634	A	2836	2	384	KTLPRTLLDILADGTILKVGVGCSEDASKLLQ
285 1635 A 2843 20 271 PIRPYYSYSGLDRDCSWLPLAKAWLPDVMIL VCDRVSEDGINRQQAQEWCIKHGFELVELSP EELPEEDGKCLCVRRKYGTYI 286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFTNSELHRAM NLHVGNLRLLSGPLDQVRAALPTPALSPKDK AVLQNLKRILAKVQEMRDQRVSLEQQLRELI QKDDITGSLVITDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE EPGQELAETALHLAVRTADQTSLHLVE DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK TDFEGQPPIPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQQQLYKVFLHGSQQQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF				1 1			DYGLVVRGCLDLRYLAMRQRNNLLCNGLSL
KR				1			
285 1635 A 2843 20 271 PIRPYYSYSGLDRDCSWLPLAKAWLPDVMIL VCDRVSEDGINRQQAQEWCIKHGFELVELSP EELPEEDGKCLCVRRKYGTYI 286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFTNSELHRAM NLHVGNLRILLSGPLDQVRAALPTPALSPKDK AVLQNLKRILAKVQEMPQRVSLEQQLRELI QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPTTRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAT SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF		1		ł			QVIYAARDAQISVALFLHLLGYPFSRNSPGEK
286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVREWAKYMEVHEKASFINSELHRAM NLHVGNLRLLSGPLDQVRAALPTPALSPKDK AVLQNLKRILAKVQEMRDQRVSLEQQLRELI QKDDITGSLVTIDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	205	1626		0042			
286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVREWAKYMEVHEKASFTNSELHRAM NLHVGNLRILSGPLDQVRAALPTPALSPKDK AVLQNLKRILAKVQEMRDQRVSLEQQLRELI QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSFKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGVVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK TIDFEGQPTTRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	483	1022	A	2843	20	271	
286 1636 A 2845 197 278 TAEDVLTVAYEHGVNLFDTAEVYAAGK 287 1637 A 2851 2 427 FVAEVRREWAKYMEVHEKASFINSELHRAM NLHVGNLRLLSGPLDQVRAALPTPALSPKDK AVLQNLKRILAKVQEMRDQRVSLEQQLRELI QKDDITGSLVTIDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK TIDFEGQPTPTRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF		į	l		İ		
287 1637 A 2851 2 427 FVAEVREWAKYMEVHEKASFTNSELHRAM NLHVGNLRLLSGPLDQVRAALPTPALSPKDK AVLQNLKRILAKVQEMRDQRVSLEQQLRELI QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNIGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	286	1626		12845	107	270	
NLHVGNLRLLSGPLDQVRAALPTPALSPKDK AVLQNLKRILAK VQEMRDQR VSLEQQLRELI QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT EVYLEQNLAAQDRVLCALT FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAK YVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETAHLAVRTADQTSLHLVE EPGGELAETAHLAVRTADQTSLHLVE DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAT SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF							
AVLQNLKRILAKVQEMRDQRVSLEQQLRELI QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNIGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	201	103/	Λ.	2031	-	421	
QKDDITGSLVTTDHSQMKKLFEEQLKKYDQL KVYLEQNLAAQDRVLCALT 288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE EPGQELAETALHLAVRTADQTSLHLVE DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF					Į.	İ	
288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF				j [. [OKDDITCEL VETDUSONAVY ESSAY VYZZA
288 1638 A 2859 2 469 FVNLGILTCIECSGIHREMGAHISRIQSLELDK LGTSELLPAKNVGNNSFNDIMEANLPSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWYYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF							
LGTSELLPAKNYGNNSFNDIMEANI,PSPSPKP TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSOFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK TIDFEGQPPTPRDKLSCWYYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	288	1638	À	2850	2	469	
TPSSDMTVRKEYITAKYVDHRFSRKTCSTSSA KLNELLEAIKSRDLLALIQVYAEGVELMEPLL EPGQELAETALHLAVRTADQTSLHLVE 289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWYYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF				2007	~	707	
289 1639 A 2861 2 454 FVASGGPATARMSDSOFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK TIDFEGQPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF					İ	ļ	
289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF							
289 1639 A 2861 2 454 FVASGGPATARMSDSQFFCVAEERSGHCAVV DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPIPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF					ſ	·	
DGNFLYVWGGYVSIEDNEVYLPNDEIWTYDI DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	289	1639	A	2861	2	454	
DSGLWRMHLMEGELPASMSGSCGACINGKL YIFGGYDDKGYSNRLYFVNLRTRDETYIWEK ITDFEGQPPTPRDKLSCWVYKDRLIYFG 290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF				ا ا	-		
290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF		[† [ſ	ĺ	
290 1640 A 2868 1 378 FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF							
290 1640 A 2868 1 378 . FRQGQLYKVFLHGSQGQVYHSQQVGPPGSAI SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF					1	1	
SPDLLLDSSGSHLYVLTAHQVDRIPVAACPQF	290	1640	Ā	2868	1	378	
	·		i		1	'	
]	j		PDCASCLQAQDPLCGWCVLQGRCTRKGQCG

SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	1 4-4
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide	""	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	
- '	uence	[M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1		}	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	İ		i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l	ł	peptide		/=possible nucleotide deletion, \=possible
			L	sequence	1	nucleotide insertion
1	 	1	ł			RAGQLNQWLWSYEEDSHCLHIQSLLPGHHPR
L	<u> </u>			l		QE
291	1641	Α	2870	1	385	FRYMPNNRQQLLRKRHIGNDIVTIVFQEPGAL
1	ļ	1	ļ	ļ	l	PFTPKSIRSHFQHVFVIVKVHNPCTENVCYSV
1			1	•		GVSRSKDVPPFGPPIPKGVTFPKSAVFRDFLL
						AKVINAENAAHKSEKFRAMATRTRQEYLKD
1 .]	j		LA
292	1642	A	2877	3	188	RPTRPPPATTQSPESTMDTSLKKEKSAILDLYI
	10.5		20,,	١	1 ***	PPPPAVPYSPRYVAVHCHGMLVSCWCHL
293	1643	A	2878	1	427	
1273	1043	l ^	2070	1	421	REKEEEVEEEEDKVVKETEKEAEQEKEEDSL
		1]	GAGTHPDAAIPSGERTCGSEGSRSVLDLVNYF
j						LSPEKLTAENRYYCESCASLQDAEKVVELSQ
1			[GPCYLILTLLRFSFDLRTMRRRKILDDVSIPLL
						LRLPLAGGRGQAYDL
294	1644	A	2879	109	245	QLCCFCFRQTTLIVYILSFIGMVIFTFTLDLRYI
L						IIVFVTGGVLG
295	1645	Α	2880	3	320	LASSQHGILNNLSLLFSICKTCIRTMDHHCPRA
{ [NNCVGEQNHRFFCALHCKSKHFCIEFTLNTNF
						FNCFLPGAEKSTIDAPFSLQPFLQDSKYNTALS
į l			i			LSESISO
296	1646	A	2892	209	363	SQYSHSLDYHLLQVTKNPFTLGDSSNPGQTE
				205		RLQEFSQKMDQVRGHWPVST
297	1647	A	2893	8	424	SPXTLXLDTFILLGIQDNILVLILATPPFMAGG
/	2047	Λ.	2073	0	724	STATEALDITILEGIQUNILVEILA IPPEMAGG
						KLYSTMGRFLRDRKNPACREMAVVLLANLA
				·		QGDSLAARAIAVQKGSIGHLLGFLEDSLAAT
]]						QIQQSQASLLHMHNPPFEPTSVDMMRRACRA
298	1648	<u>A</u>	0004		445	LLALAKVDDNHSEF
298	1046	A	2894	310	445	FWIYFPSFFMTGYLPLGFEFAVEITYPESEGTS
	1.618					SGLLNASAQVNL
299	1649	Α	2898	1	492	KIKAKNLTNYDLCSIFLGTSTLLVWVGVIRYL
1	l					GYFQAYNVLILTMQASLPKVLRFCACAGMIY
1 1	1					LGYTFCGWIVLGPYHDKFENLNTVAECLFSL
				ĺ		VNGDDMFATFAQIQQKSILVWLFSRLYLYSFI
1						SLFIYMILSLFIALITDSYDTIKKFQQNGFPETD
1						LQEF
300	1650	A	2901	1	445	PVWWNSLNGASEVTFSVHVKDGGSFPKTDST
i j	}]		J	TVTVRFVNKADFPKVRAKEQTFMFPENQPVS
1						SLVTTITGSSLRGEPMSYYIASGNLGNTFQIDQ
					}	LTGQVSISQPLDFEKIQKYVVWIEARDGGVPP
	ļ	1		}		FSSYEKLDITVLDVNDNAPIF
301	1651	A	2902	162	433	THFICLPLGYCFPLLDKDLOLPSGFNCNFDFLE
[·	I				EPCGWMYDHAKWLRTTWASSSSPNDRTFPG
		l	į			KPAVSEDMKELRPACSTYFNPRFPYKL
302	1652	A	2909	2	412	GPQMLCKKIYFIWVTRSQCQFEWLADIMQEV
				~	714	
ŀ	- 1		1		ĺ	EENDHQDLVSVHIYVTQLAEKFDLRTTMLYI
	ľ	Ì	}			CERHFOKVLNRSLFTGLRSITHFGRPPFEPFFN
	l			-	1	SLQEVHPQVRKIGVFSCGPPGMTKNVEKACQ
202	1652					LVNRQDRAHFM
303	1653	A	2914	291	453	KLNRWLCFFYSWSFGILLYEMVTLGAPPYPE
	1660		4000			VPPTSILEHLQRRKIMKRPSSCS
304	1654	A.	2926	179	354	PGVPSQALRKAESLKKCLSVMEAKVKAQTAP
		[NKDVQREIADLGEVGAASLPPSSGPGA
305	1655	A	2938	135	438	GMGYLHAKGILHKDLKSKNVFYDNGKVVIT
	ļ	l	ļ	ļ	1	DFGLFSISGVLQAGRREDKLRIQNGWLCHLA
	1	ļ	- 1	j	1	PEIIRQLSPDTEEDKLPFSKHSDVFALGTIWYE
				1	ł	LHAREWP
306	1656	A	2944	2	329	VRWNSCVNCSCAFGNGASLSTSLGESSGCLW
	- 1	1		1		EIGKWLSCSLLSFPSPLAVLIITFCIVTVLGREA
[ſ	1	[ĺ	1	LTKGALWAVFLLAGSALLCAEVTGVIWROPE
						21120 DWATT DEMOCRATION TO TWEETER

Second 1979	070 10	T éro m	11/-4	Toro	T 70 - 11 - 1 - 1	1 5 2	
mucleotide muc	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence		1	noa				
Seq			1				
uence 914 ng to first amino seid residue of peptide peptide requence peptide sequence			1	1			
minto seid recidue of peptide requence page p		иенсе	[
residue of peptide sequence P-Tyryosine, X-Uaknown, *-Siop codon, P-possible nuclotide deletion, t-possible nuclotide insertion	uence			714			
peptide pequence peptide peptide peptide percent per			ļ				
	l	ţ	İ	1		sequence	
SKYKLSPKVSSA	}			Ì		}	
1657	<u> </u>	 	 -		sequence	<u> </u>	
POLSTITEGSHAFLPCKARGSPEPNITWOKDGO PYSGAEGKFTIPOSEGLIVALNEGODAGTYT CTAENAVGRARRAVHLTILVLPYFTTLPGDRS LBLGDRJWLR	207	1657	-	2050		411	I
Pysgaeckftipgsgellvrnlegopagtyt	307	1037	^	2930	2	411	
CTAENAYGRARRAVBILTILVLPVFTTLPGDRS	ļ	1	1			J	
LRLGDRLWLR		1					CLVENTANCE VEDENTI LII ALEACTI DODGO
1658				ļ	1	!	
DSSLPEEEEDEDKAINGSGMAENRERRISES DWMKTYPSYNQTNSSMDFRNYMRDETLEP LPKNWEMAYTDTGMIYPIDHYNMRDETLEP LPKNWEMAYTDTGMIYPIDHYNMRDETLEP LPKNWEMAYTDTGMIYPIDHYNMRTUTTULDP RLCKKAKAPEDC DPLITLITEPTGLLYYGAREALFAFSMEALE LQGAVRGGAVGGSRACQRARPRGAVLG DQMMRAIIDTFYGHDVVPEOSYVQMFPYPC	308	1658	Δ	2051	1	407	
DWMKTVPSYNQINSSMDFRNYMMRDETILDP PKNWEMAYTDTGMYPIDHNYKTTTWLDP RLCKKAKAKPEDCFFERWAYSVAMTO RLCKKAKAKPEDCFFERWAYSVAMTO RLCKKAKAKPEDCFFERWAYSVAMTO RLCKAYAWATMECF RLCKYYAWATMECF RLCKYYAWATMECF RLCKYYAWATMECF RLCKYYAWATMECF RLCKYYAWATMECF RLCKYYAWATMECF RLCKYYAWATMECF RLCKYYAWATMECF RLCKYYAWATMECF RLCKYYAWATMECF RLCKYYAWATMECF RLCKYYAWAYATMECF RLCKYYAWATMECF RLCKYYAWATMECF RLCKYYAWATMECF RLCKYYAWATMECF RLCKYYAWATMECF	1 300	1030	^	2,5,1	1 *	1 707	
LPKNWEMA YIDTGMIYFIDHNIKTTIWLDP RLCKKAKAPDC						ļ	
RICKKAKAPEDC]			j	İ	1	
1659			1		i		
LQGAVRGGAVGGSRACQRARPRGAVLG	300	1659	Ι	2954	2	170	
1660	307	1057	\ \frac{1}{2}	2754		117	
	310	1660	Α	2050	1	410	
HIVAEKEHRILKEVMKTMGLNNAVHWVAWFI TGFVQLSISVTALTALIKYGQVLMHSHVVIIW LFLAVYAVATIMFCF] 310	1000	^	2737	*	419	
1661	ì						
	1		1	ł			1
1661	ļ			ł		ł	I FLAVVAVATIMECE
PWVPHLIPPSSPGYLGVMKAQKPGAGEGMK PQKPGLRGTLKPQKSGHIGHERGPWPGPCNA RVAPMLIPPLTPTPOVPDKEGGWGLKSQPPS AVQNGKLPGHPPNGYGPGAEPGFNGGLEPQ KI	311	1661	A	2963	3	465	
POKPGLRGTLKPOKSGHGHENGPWPGPCNA		100.	**	2,03	۱	103	
RVAPMLLPRLPTTGVPSDKEGGWGLKSQPPS	1		l				
AVQNGKLPGHQPPNGYGPGAEPGFNGGLEPQ XI	l						
1662	1	1				1	
1662	ŀ	l	1	ł			
HTGGCSPKTKPHIKEECIVPTPCYKPKEKLPV EAKLPWFKQAQELEEGAAVSEEPSFIPEAWS ACTVTCGVGQVEGAVEGAVSEEPSFIPEAWS ACTVTCGVGQVEVROCQVILSFSQSVADLPI DECEGPKPA 313 1663 A 2969 2 430 VVADNCRQGYLDALFILERRGLTKEPVLWT LVSKEPPAPADGNWDAGCDQRRKGGLSLNW KVPHVQVKDVPNFEQLSPELEAALKKACTRD PSRWARFWHSGPGQVLTYLLLPCTLPFEYIYF RSRRLVVWLPDVPADLWWMQ 314 1664 A 2971 422 33 LDXSHNALQRLRPGWLAPLFQLRALHLDHNE LDALGRGVFVNASGLRLLDLSSNTLRALGRH DLDGLGALEKLLLFNNRLVHLDEHAPHGLRA LSHLYLGCHELASFSFDHLHGLSATHLLTLDL SSNRM 315 1665 A 2973 1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGFVTILQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT 316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFITYADLGALTKIRRHDNTGNR AGWFLDRIDITDMNNEITYYPPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQVDASSNLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSLKYTKCLFTEQFKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTTPGECHVAAVDREVAAGPPVPAEGISGETTH	312	1662	A	2967	3	405	
SAKLPWFKQAQELEEGAAVSEEPSFIPEAWS ACTVTCGVGTQVRIVRCQVLLSFSQSVADLPI					_		
313 1663 A 2969 2 430 VVADNCRQGYLDALRFLERRGLTKEPVLWT LVSKEPPAPADGNWDAGCDQRRKGGLSLNW KVPHVQVKDVPNFEQLSPELEAALKKACTRD PSRWARFWHSGPGQVLTYLLLPCTLPFEYIYF RSRRLVWHPDVPADLWWMQ 314 1664 A 2971 422 33 LDXSHNALQRLRPGWLAPLFQLRALHLDHNE LDALGRGVFVNASGLRLLDLSSNTLRALGRH DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM 315 1665 A 2973 1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT 316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGGTDTFTTYAIDLGALTKIRRHEDNTGNR AGWFLDRIDITDMNNEITYYPPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDEWMQVNSNYBAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTIM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIRKSSKVTKCLFTEQPKPHILRFA ENYDARLLRIDIANTLREQVQCLFNKTYGKQ RTITGEGHVAAVDREVAGFFVPAEGISGETH			l	ł			
DECEGPKPA			İ				
LVSKEPPAÄDGNWDAGCDQRRKGGLSLNW KVPHVQVKDVPNFEQLSFELEAALKKACTRD PSRWARFWHSGPGQVLTYLLLPCTLPFEYIYF RSRRLVYWLPDVPADLWWMQ 314 1664 A 2971 422 33 LDXSHNALQRLRPGWLAPLFQLRALHLDHNE LDALGRGVFVNASGLRLLDLSSNTLRALGRH DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM 315 1665 A 2973 1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPFGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT 316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYADLGALTKIRRHDNTGNR AGWFLDRIDITDMNNEITYYPPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPHIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGIVAAVDREVAGFPVPAEGISGETH			•		•		
LVSKEPPAPADGNWDAGCDQRRKGGLSLNW KVPHVQVKDVPNFEQLSFELAALKKACTRD PSRWARFWHSGPGQVLTYLLLPCTLPFEYIYF RSRRLVVWLPDVPADLWWMQ 314 1664 A 2971 422 33 LDXSHNALQRLRPGWLAPLFQLRALHLDHNE LDALGRGGVFVNASGLRLLDLSSNTLRALGRH DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM 315 1665 A 2973 1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPFGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT 316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYADLGALTKIRHDNTGNR AGWFLDRIDITDMNNEITYYPPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTIM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPHILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGIVAAVDREVAGFPVPAEGISGETH	313	1663	Α	2969	2	430	
KVPHVQVKDVPNFEQLSPELEAALKKACTRD PSRWARFWHSGPGQVLTYLLLPCTLPFEYIYF RSRRLVVWLPPUPADLWWMQ 314 1664 A 2971 422 33 LDXSHNALQRLRPGWLAPLFQLRALHLDHNE LDALGRGVFVNASGLRLLDLSSNTLRALGRH DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM 315 1665 A 2973 1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTILQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT 316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGOTDTFTTYAIDLGALTKIRHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPVVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEQQ 318 1668 A 2995 119 414 LPEKEFPIIKSSSLKVTKCLFTEQPKPIILRFA ENVDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETTH	1						
RSRRLVVWLPDVPADLWWMQ 314 1664 A 2971 422 33 LDXSHNALQRLRPGWLAPLFQLRALHLDHNE LDALGRGVFVNASGLRLDLSSNTLRALGRH DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM 315 1665 A 2973 1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRISEDPTFLAVAENOPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT 316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTTYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYPPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENVDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETTH		ļ	Ì				
1664							PSRWARFWHSGPGQVLTYLLLPCTLPFEYIYF
LDALGRGVFVNASGLRILDLSSNTLRALGRH DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM 315 1665 A 2973 1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT 316 316 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTTYGEEYGDTGERPLKKSDKSNK FEQGQTDTFITYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYPPCQRWLAVEE DDGQLSRE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH							RSRRLVVWLPDVPADLWWMQ
DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM 315 1665 A 2973 1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT 316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTIYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYPPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQVDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ TARKVIVTVQARPAVPMCWTEGQ RTPGEGHVAAVDREVAGFPVPAEGISGETIH	314	1664	Α	2971	422	33	LDXSHNALQRLRPGWLAPLFQLRALHLDHNE
LSHLYLGCNELASFSFDHLHGLSATHLLTLDL SSNRM 315 1665 A 2973 1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT 316 1666 A 2978 2 400 ELVVELVSAGKSGPBRNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH		,					LDALGRGVFVNASGLRLLDLSSNTLRALGRH
SSNRM SSNRMSPGRLTYSLQLSEDSKAFRIHPQ TGEVTILQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT SSNTGTVIIVAVLDLNDNT SSNTGTVIIVAVLDLNDNT SSTTGTVIIVAVLDLNDNT SSTTGTVIIVAVLDLNDNT STTGTVIIVAVLDLNDNT STTGTVIIVAVLDLNDNT STTGTVIIVAVLDLNDNT STTGTVIIVAVLDLNDNT STTGTVIIVAVLDLNDNT STTGTVIIVAVLDLNDNT STTGTVIIVAVLDLNDNT SSNTGTVIIVAVLDLNDNT ST	'	j	J]			DLDGLGALEKLLLFNNRLVHLDEHAFHGLRA
1665 A 2973 1 525 ITVSTHASGSPFGLEPQSGWLWVRAALDREA QELYILKVMAVSGSKAELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT	İ		1				LSHLYLGCNELASFSFDHLHGLSATHLLTLDL
QELYILKVMAVSGSKÄELGQQTGTATVRVSI LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT 316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYAIDLGALTKIRIRHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH	L		<u> </u>	<u> </u>			SSNRM
LNQNEHSPRLSEDPTFLAVAENQPPGTSVGRV FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT 316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVFFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKYTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH	315	1665	A	2973	1	525	
FATDRDSGPNGRLTYSLQQLSEDSKAFRIHPQ TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT 316 316 316 317 318 318 318 318 318 318 318 318 318 318			Ι.	1			
TGEVTTLQTLDREQQSSYQLLVQVQDGGSPP RSTTGTVIIVAVLDLNDNT 316 1666 A 2978 2 400 ELVVELVSAGKSGPBRNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTIYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH		ł	ł	1			
RSTTGTVIIVAVLDLNDNT 316 1666 A 2978 2 400 ELVVELVSAGKSGPERNTYEVQVVTGNVPKA GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTIYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH		Į.	l	1			
316			1	1			
GTDANVYLTIYGEEYGDTGERPLKKSDKSNK FEQGQTDTFTTYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH			<u> </u>				
FEQGQTDTFTTYAIDLGALTKIRIHDNTGNR AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH	316	1666	j A	2978	2	400	• • • • • • • • • • • • • • • • • • • •
AGWFLDRIDITDMNNEITYYFPCQRWLAVEE DDGQLSRE 317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIRKSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH			1				**
317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH			l				
317 1667 A 2981 3 440 VLNCQGRPTRPVRINGDGQEVLYLAESDNVR LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH	1		ł	Į.	-		· · · · · · · · · · · · · · · · · · ·
LGCPYVLDPDDYGPNGLDIEWMQVNSNPAH HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH		1255	<u> </u>	200			
HRENVFLSYQDKRINHGSLPHLQHRVRFAAS DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH	317	1667	A	2981	3	440	
DPSQYDASINLMNLQVSDTATYECRVKKTTM ATRKVIVTVQARPAVPMCWTEGQ 318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH		ł	ł	1			
318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH		}	ĺ				
318 1668 A 2995 119 414 LPEKEFPIIRKSSSLKVTKCLFTEQPKPIIILRFA ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH			1	İ		•	
ENYDARLLRIDIANTLREQVQELFNKTYGKQ RRTPGEGHVAAVDREVAGFPVPAEGISGETIH	-	1.772	<u> </u>	2007			
RRTPGEGHVAAVDREVAĞFPVPAEGISGETÜH	318	1668	A	2995	119	414	- 1
		1	1]			
1 1009 A 2999 Z 332 GFFAYTYGRLVVVEDLHSGAOOHWSGHSAFI	210	1	 	10000			
- John Marie Comment of the Comment	319	1009	A	2999	2	332	GFFAYTYGRLVVVEDLHSGAQQHWSGHSAEI

SEQ ID	SEQ ID	Met	SEQ	Predicted	I Designation of	
NO: of	NO: of	hod	ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1 1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	Ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	l]	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
}		ì	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion. \=possible
		<u> </u>	<u></u>	sequence		nucleotide insertion
	l	l	}	l		STLALSHSAQVLASASGRSSTTAHCQIRVWD
]	ĺ				VSGGLCQHLIFPHSTTVLALAFSPDDRLLVTL
320	1670	A	3000	693	222	GDHDGRTLALWGTGHL
320	1070	^	3000	023	322	IDESTGLIITVNYLDYETKTSYMMNVSATDQA PPFNQGFCSVYITLLNELDEAVQFSNASYEAA
ł	ł	l		}	l	ILENLALGTEIVRVQAYSIDNLNQITYRFDAY
						TSTQAKALFKIDAITVRGWGQGAPFFPI
321	1671	A	3001	6	383	RIPRGKACXTVLGRSTGELEGFASSRLPPQPC
						GWGQSSDLLSRIDLDELMKKDEPPLDFPDTLE
1		l				GFEYAFNEKGQLRHIKTGEPFVFNYREHLHR
		ļ				WNQKRYEALGEIITKYVYELLEKDCNSKKVS
322	1672	A	3007	192	447	ERVRNSLFPGRGDSQCACCPSSPVWVFLETGF
		ļ				LFPWLFLQVEVIKKAYMQGEVEFEDGENGK
						DGAASPRNVGHNIYILAHQLARH
323	1673	A	3019	18	245	KELLFYHLIVNNINFFNTRYAKIHIPILASVSEH
						QPTTWVSFFFDLHILVCTFPAGLWFCIKNIND
324	1674	A	3020	500		ERVFGKRGF
324	10/4	A	3020	523	797 .	LCYFSARYHQRKIFGILYIFTLSAINRKEPNLFI
						YLFIFFEMESHSVTHAGVQRHNLNSLQPLPPG
325	1675	Α	3022	2	156	FKRFSCLCFLSSWNYRGAPPGPANF NDFLPLYFGWVLTKKSSETLRKAGQVFLEEL
323	10/3	**	3022	4	150	GNHKAFKKELRQCRWQVGAL
326	1676	A	3023	38	172	KMVRGSKKLISFFPGGPYGILAGRDPSKGLAT
				•		FCLNKEALKDEFE
327	1677	A	3027	1 .	385	LTLEFLLLPAASELAHGKRLACCIVDHKLPEC
						GFYGLYDKILLFKHDPTSANLLQLVRSSGDIQ
						EGDLVEVVLSASATFEDFOIRPHALTVHSYRA
			I	•		PAFCDHCGEMLFGLVRQGLKCDGCGLNYHK
200	1650					RC
328	1678	Α	3030	13	569	ITRPTISCQRPGPGLAAGMLPYTVNFKVSART
ļ					,	LTGALNAHNKAAVDWGWQGLIAYGCHSLV
						VVIDSITAQTLQVLEKHKADVVKVKWAREN YHHNIGSPYCLRLASADVNGKIIVWDVAAGV
						AQCEIQEHAKPIQDVQWLWNQDASRDLLLAI
						HPPNYIVLWNADTGTKLWKKSYADNILSFSF
}						D
329	1679	A	3038	90	744	SVNLPPSLWPWEEAMDSTKSEPLKGSPEAED
			1	-		GNIEYKKLVNPSQYRFEHLVTOMKWRLOEG
			1	.		RGEAVYQIGVEDNGLLVGLAEEEMRASLKTL
}		J	J			HRMAEKVGADITVLREREVDYDSDMPRKITE
İ						VLVRKVPDNQQFLDLRVAVLGNVDSGKSTL
		1		Ì		LGVLTQGELDNGRGRARLNLFRHLHEIQSGR
1					Ì	TSSISFEILGFNSKGEVHGINGTQWGQTLRMG
330	1680	A	3040	3	207	W
230	1000	^	JU4U	,	397	LCSTLLLLTIPSWVLSQITLKESGPTLMKPTET
ļ	1	ŀ				LTLTCTFSGFSLNTSGVGVAWIRQPPGKALE WLALIYWDDDKRYSPSLNDRLTIAKDTSRNQ
l	}	.				VVLTMTNMGPVDTATYYCAQFARGARGSN
ŀ]	ŀ			,	WFDPWGQ
331	1681	A	3043	3	1509	AGIRHEAPPTTSNRHRRQIDRGVTHLNISGLK
	•	i			• •	MPRGIAIDWVAGNVYWTDSGRDVIEVAQMK
	1	- 1				GENRKTLISGMIDEPHAIVVDPLRGTMYWSD
		1	1			WGNHPKIETAAMDGTLRETLVQDNIQWPTG
1		- 1	İ	i		LAVDYHNERLYWADAKLSVIGSIRLNGTDPI
1				l	İ	VAADSKRGLSHPFSIDVFEDYIYGVTYINNRV
	l	ľ		ļ		PKIHKFGHSPLVNLTGGLSHASDVVLYHQHK
1				ŀ	ļ	QPEVTNPCDRKKCEWLCLLSPSGPVCTCPNG
- 1	- 1	ľ	- 1	!	ł	KRLDNGTCVPVPSPTPPPDAPRPGTCNLQCFN
	1		1			GGSCFLNARRQPKCRCQPRYTGDKCELDQC

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A-Alanine C=Cysteine, , D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion WEHCRNGGTCAASPSGMPTCRCPTGFTGPKC TQQVCAGYCANNSTCTVNQGNQPQCRCLPG FLGDRCQYRQCSGYCENFGTCQMAADGSRQ
						CRCTAYFEGSRCEVNKCSRCLEGACVVNKQS GDVTCNCTDGRVAPSCLTCVGHCSNGGSCT MNSKMMPECQCPPHMTGPRCEEHVFSQQP GHIASILIP
332	1682	A	3045	3	952	TTTISNFHTQVNRTYCCGTYRAGPMRQISLVG AVDEEVGDYFPEFLDMLEESPFLKMTLPWGT LSSLRLQCRSQSDDGPIMWVRPGEQMIPTAD MPKSPFKRRSMNEIKNLQYLPRTSEPREVLF EDRTRAHADHVGQGFDWQSTAAVGVLKAV QFGEWSDQPRITKDVICFHAEDFTDVVQRLQ LDLHEPPVSQCVQWVDEAKLNQMRREGIRY ARIQLCDNDIYFIPRNVIHQFKTVSAVCSLAW HIRLKQYHPVVEATQNTESNSNMDCGLTGKR ELEVDSQCVRIKTESEEACTEIQLLTTASSSFP PASE
333	1683	A	3046	497	167	SACSTGPELPGRATRSLTRPANQKGCDGDRL YYDGCAMIAMNGSVFAQGSQFSLDDVEVLT ATLDLEDVRSYRAEISSRNLAVSAPVDTCVG CSSKTWKVAPFVRAWWRP
334	1684	A	3053	37	276	VITDLEEQLNQLTEDNAELNNQNFYLSKQLD EASGANDEIVQLRSEVDHLRREITEREMQLTS QKQVRRVNKVVRSLEDF
335	1685	A	3054	2	846	WDAWGDWSDCSRTCGGGASYSLRRCLTGR NCEGQNIRYKTCSNHDCPPDAEDFRAQQCSA YNDVQYQGHYYEWLPRYNDPAAPCALKCH AQGQNLVVELAPKVLDGTRCNTDSLDMCISG ICQAVGCDRQLGSNAKEDNCGVCAGDGSTC RLVRGQSKSIIVSPEKREENVIAVPLGSRSVRI TVKGPAHLFIESKTLQGSKGEHSFNSPGVFVV ENTTVEFQRGSERQTFKIPGPLMADFIFKTRY TAAKDSVVQFFFYQPISHQWRQTDFFPCTVT CGGG
336	1686	A	3058	54	347	VVGKQEAGAHSDSCCLLHTPPRLTPAHSRKA LRNSRIVSQKDDVHVCIMCLRAIMNYQVSRG AWDWRLGSPACPHWGLHKLPRLWDPLSLYP VLCWGT
337	1687	A	3059	2	709	ILTSLVELTRFETLTPRFSATVPPCWVEVQQE QQQRRHPQHLHQQHHGDAAQHTRTWKLQT DSNSWDEHVFELVLPKACMVGHVDFKFVLN SNITNIPQIQVTLLKNKAPGLGKVNGLRLCPF LEDHKEDILCGPVWLASGLDLSGHAGMLTLT SPKLVKGMAGGKYRSFLIHVKAVNERGTEEI CNGGMRPVVRLPSLKHQSNKGYSLASLLAK VAAGKEKSSNVKNENTSGTRK
338	1688	A	3060	85	384	KAFYNYHVLELLQMLVTGGVSSQLEQHLDK DKVYGVADSCTSLLSGRNRCKLGI.LSLHETIL SDVNPRNTFGQLFCGSLDLFGILCVGLYRIIDE EELNP
339	1689,	A	3063	236	362	CFLCLSGDFMVMTIFFNVSRRFGYVAFQNYV PSSVTTMLSWV
340	1690	A .	3065	3	1249	DLWQFTPLHFAASKNRVEVCSLLLSYGADPT LLNCHNKSAIDLAPTPQLKERLAYEFKGHSLL QAAREADVTRIKKHLSLEMVNFKHPQTHETA LHCAAASPYPKRKQICELLLRKGANINEKTKE FLTPLHVASEKAHNDVVEVVVKHEAKVNAL DNLGQTSLHRAAYCGHLQTCRLLLSYGCDPN

NO: of nucl- peptide eotide sequence uence nucle to the peptide sequence nucle to the peptide sequence nucle to the peptide sequence nucle to the peptide in nucleotide location corresponding to last amino acid residue nucleotide location corresponding nucleotide location to last amino acid residue nucleotide location corresponding nucleotide location corresponding nucleotide location corresponding nucleotide location corresponding nucleotide location nucleotide location nucleotide location nucleotide nucleotid	lycine, H=Histidine,
nucl- peptide in nucleotide location F=Phenylalanine, G=G USSN location corresponding uence 09/496 correspondi to last amino M=Mcthionine, N=Asguence 914 ng to first acid residue Q=Glutamine, R=Argin	lycine, H=Histidine,
eotide sequence USSN location corresponding le-Isoleucine, K=Lysine corresponding uence 09/496 corresponding to last amino M=Mcthionine, N=Asguence 914 ng to first acid residue Q=Glutamine, R=Argin	
sequence under under 914 ng to first acid residue Q=Glutamine, R=Argin	
uence 914 ng to first acid residue Q=Glutamine, R=Argin	
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	nine C=Carina
amino acid of peptide T=Threonine, V=Valin	w. W-Tamtonhon
residue of sequence Y=Tyrosine, X=Unkno	c, w-rryptopnan,
peptide /=possible nucleotide d	olotion _moneille
sequence nucleotide insertion	eletion, —possible
	INTSSISITALAAEIKNPER
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	LVEVVSGLATAAEKNPEK
OI CELTI CHICKONI	RCHSTPGFIVNRVARPY
353 1703 A 3111 3 188 HFSLFRIAFAVFLTY	APEVI
The last to the two transfers to the tra	MTVGLPLPVIPLFVHHEL
354 1704 A 3116 367 225 WOLFILNGTEING	FLATVLTRGYAGRLA
1 1 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -	ETDTESCVNGWYYDRSS
FPFSNMTEVRGLVFI	<u>.S</u>
355 1705 A 3117 101 53 VINLVYLISSPRPELK	PVDKESEVVMKFPDGF
EKFSPPILQLDEVDF	YYDPKHVIFSRLSVSADL
ESRICVVGENGAGKS	STMLKLLLGDL\APVRGI
RHAHRNLKIGYFSQI	HHVGAAGT*TFSACGNL
LGTQVFLGRPEEEYV	RHQLGFGMGISGELGHA
SSLPACLGGQKEAEV	VAFCSDGLLPCPNFL\IL\
DEPTNHLGHGRAIE	ALGPCLQTISGVGVILVS
HE*SALSRLVCRELLY	
356 1706 A 3121 137 466 RGGRDWGEHNQRLE	EEHQARAWQGAMDAG
AASREHARWOGTGI	APGTRVAVAPTCVOGL
POERSVCRPFFSSRW	REGPVWALGAGAHGKP
RWSGGVRCVVRGGF	RWFTPAPH
357 1707 A 3124 1249 229 MLEAPGPSDGCELSN	NPSASRVSCAGQMLEVQ
	DHLREAGITAVLTVDSE
EPSEK A GPG V EDI WI	RLFVPALDKPETDLLSH
LDRCVAFIGOARAFO	GRAVLVHCHAGVSRSV
	EKAYEKLOILKPEAKMN
FGEFWOI KI YOAMO	GYEVDTSSAIYKQYRLQ
KVTEK YPEI ONI POE	ELFAVDPTTVSQGLKDE
VI VY CRY CDD CI EDG	SSILDHREGSGPIAFAH
KPMTPSSMI TTGPO	AQCTSYFIEPVQWMESA
LI CVMDCOLL CDVC	SAVI GSENIUMGEOGG
COUNTRACOUNTRY	SAKLGSFNWYGEQCSC DEMKILPVLGSQTGKI
358 1708 A 3127 816 139 EVETLGPRTPGP/EAC	DEMKILP VLGSQTGKI
DVEIDGIRII GIVER	SPTPGSCPGWQEPSPGP
	/LGKLLPDPEETPAGKTP
	ANFSPGAAA*FGGALSPP
GODL/GHMLLQGPPS	PFRLQQQ*QTPPGSHSP
	DTRSCWGHKRSWRGW
	P*PAPAGIP/GRPTWEGG
	PVWRGKRGSANGFLSW
359 1709 A 3132 3 191 HEHLLLILLCVFLVK	
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	SQGVNDNEEGFFSARG
	PALADVITYCDYRAQIA
*AASTPKRAASIAHN	
	LCCPGWNAVVQSQLT
	SSWDYRHVPEYPANFL
*RQGFPMLPRLVSNS	WAQTVHPPRPPKVLDL
QA	1
361 1711 A 3135 56 1449 PVPAPRVSPSARGAP	GRPRLPGVRGPRHS/WA
AD*RGSRM/PPRAPAI	PSPTGP/APGGKKVRGR
VPEDPDAYEPRCSAL	*V*PTHVTSPQFCDP*N
GOIRSYFTVLLRGLNI	ETMLVK/PLCRREP/PEA
	CHEDPRGAGROWDAD
	PGRHMWMRLCLAAQQ
	RLTEPEAWARRHRRPW
1 1 1 1 1 1	PPPSHQGRRTNTDPSAT
	APASGPRGWRRGMPO
	SHSPRPREAPLRAIHPA
	VIYGWVTLFTPPEAGT
	PPVPPTQMGLRISGLPR
	QLAFQCHLPHDEVGPP
	ATU ACUTU UDEACLE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						RNQSPLONDTLSSGLPMGPRRQVWPLARVG GHSSPREPQVLKKPLWGQTDIAGVGSASLYP DNL
362	1712	A	3136	1270	274	RVGMVLGTREVGDSTPPPSPPLYPFTGNEFVQ HNTWQLSRVYPSDLRTDSSNYNPQELWNAG CQM/V*GGSRDWEEGVEEQQVGNKFSSDGR VGECSRKLLG*EMLSVDITSRYRAPSTYLLNS LKEGLEGLHGESCSSFLLGPSVAMNMQTAGL EMDICDGHFRQNGGCGYVLKPDFLRDIQSSF HPEKPISPFKAQTLLNQVISVQQLPKVDKTKE GSIVDPLVKVQIFGVRLDTARQETNYVENNG FNPYWGQTLCFRVLGPDFPMLRFGKMDYDW KSRNDLLGKTPCPGTCMQQGYRHIHLLSKDG ISLRPASIFVYICIQEGLEGDES
363	1713	С	3139	60	248	MFAGSYGKSMFSFSKKVLNCLPKWRYHFVIA
364	1714	Α	3140	57	418	PAMNESPLAPHLHQHLVFSVFQVLTILIGV** SAFKTLQLPAFSLYFDLGSLKLLILRIHTSIVK NHKVESPRTMSPG*DPQSFLQIPQPRPPQLRV GLTSGLIQHFHSPSSCQFPLLRGPPFPRQPPLGI SGASLCPVLSPPR*PLQPSSL
365	1715	Α	3145	122	413	LLPYPSLFVFLRQCHFVTRLECNGVVSAHCN LHLPGSSDSPASAS*VAGTTGVCHHTRLIFVF LV*TGFHYVAQAGLELLTA*SVPPQLPKVVGL QA
366	1716	A	3150	247	2	VGEKLHDIRFGNDFDMTPKAQATKEKIDKLN FIKIKKLCIEGYY/NREPQNGRKIFANYVS'DK GLMATIYEELLKLSNKLIQ
367	1717	A	3152	3	2367	QKLKQNQPKRAHVEDGGSRSKQGNEQSKKT PIEKSDFAAATHPRAFYLSKPDETPNAWMSD SGTGLTYWKLEEKDMHHSLPETLEKTFISLSS TDVSPNQVLTLDPTLHMKPKQQISGIQPHGLP NALDDRISFSPDSVLEPSMSSPSDIDSFSQASN VTSQLPGFPKYPSHTKASPVDSWKNQTFQNE SRTSSTFPSVYTITSNDISVNTVDEENTVMVAS ASVSQSQLPGTANSVPECISLTSLEDPVILSKIR QNLKEKHARHIADLRAYYESEINSLKQKLEA KEISGVEDWKITNQILVDRCGQLDSALHEATS RVRTLENKNNLLEIEVNDLRERFSAASSASKI LQERIEEMRTSSKEKDNTIIRLKSRLQDLEEAF ENAYKLSDDKEAQLKQENKMFQDLLGEYES LGKEHRRVKDALNTTENKLLDAYTQISDLKR MISKLEAQVKQVEHENMLSLRHNSRIHVRPS RANTLATSDVSRRKWLIPGAEYSIFTGQPLDT QDSNVDNQLEETCSLGHRSPLEKDSSP/GSSST SLLIKKQRETSDTPIMRALKELDEGKIFKNWG TQTEKEDTSNSLL*/INPRQTETSVNASRSPEK CAQQRQKRLNSASQRSSSLPPSNRKSSTPTKR EIMLTPVTVAYSPKRSPKENLSPGFSHILLSKN ESSPIREKTYSEKATDNHVNHSSCPEPVPNGV KKVSVRTAWEKNKSVSYEQCKPVSVTPQGN DFEYTAKIRTLAETERFFDELTKEKDQIEAAL SRMPSPGGRITLQTRLNQVKCLSLNIL
368	1718	A	3163	2	2350	EFKSGGCGAGLVAAGAVLVLYPASRAGERT RVPGSPAPSSLPLHSPGACGTEVDMDPQRSPL LEVKGNIELKRPLIKAPSQLPLSGSRLKRRPDQ MEDGLEPEKKRTRGLGATTKITTSHPRVPSLT TVPQTQGQTTAQKVSKKTGPRCSTALATGLK NQKPVPAVPVQKSGTSGVPPMAGGKKPSKRP AWDLKGQLCDLNAELKRCRERTQTLDQENQ

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A-Alanine C-Cysteine,
NO: of nucl-	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-	ſ	USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ucnec	i	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
испсс	ĺ	ſ	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	İ	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
İ	İ	[peptide	sequence	/=possible nucleotide deletion, \=possible
1				sequence	[nucleotide insertion
<u> </u>	 			sequence	<u> </u>	QLQDQLRDAQQQVKALGTERTTLEGHLAKV
		}	J	l		QAQAEQGQQELKNLRACVLELEERLSTQEGL
			1	ļ		VQELQKKQVELQEERRGLMSQLEEKERRLQT
				ĺ		SEAALSSSQAEVASLRQETVAQAALLTEREER
ļ	j	J		l	ļ	LHGLEMERRRLHNQLQELKGNIRVFCRVRPV
		Ì	l			LPGEPTPPPGLLLFPSGPGGPSDPPTRLSLSRSD
			ł			ERRGTLSGAPAPPTRHDFSFDRVFPPGSGQDE
ļ	}	ļ	1	ĺ		VFEEIAMLVQSALDGYPVCIFAYGQTGSGKTF
						TMEGGPGGDPQLEGLIPRALRHLFSVAOELSG
Į.		l				QGWTYSFVASYVEIYNETVRDLLATGTRKGQ
		l	l			GGECEIRRAGPGSEELTVTNARYVPVSCEKEV
						DALLHLARQNRAVARTAQNERSSRSHSVFQL
ł						QISGEHSSRGLQCGAPLSLVDLAGSERLDPGL
		ĺ	ľ			ALGPGERERLRETQAINSSLSTLGLVIMALSN
		1	i			KESHVPYRNSKLTYLLQNSLGGSAKMLMFV
						NISPLEENVSESLNSLRFASKVEPSVLFGTAQS
		ĺ	ĺ			NRKWKTDPDLCVCVCVCVCVCVCVCVP
						MSMYRVRGGRVAGGCFIGWRAPCPRAIK
369	1719	A	3165	365	12	GYTSQGRWIDIERGPLTANTESLHENNFNALP
						GYIRKIE*I*IYKKN:*INFGGVGLLNIVKISILS/K
						IYRFDAIPVKILTRFFINLDKLILKFVLKTKIAK
						NRIKTFYIMRRKKLGDSS
370	1720	Α.	3170	393	42	GASISPSAVIDGVEGLKPMQEQEAQEAGPCLD
						*HMAPEQWVAPR\RLLFRLIFSVLHALIIAAAA
	,	1				QSSAEEDEDPRN*GQSSEDQAPNQNGLIVIVH
371	1721	A	3173	770.	510	RVHVPLGAAATVPVHRSHFPR
3/1	1/21	A	31/3	770.	210	GNGGCGLSQIPPSHLGAFSRGSLLSRG\DPRGP PPHPVIFFVFVVE\QGFTVLARMVSIS*PCDPP
						ALASQSAGITGVSHLARPONLYF
372	1722	A	3180	381	76	RVLHHDNVPAHSSPQKREISQEFQLEIRHLP*S
		**	1000			PDLAPSGCFLFLNLKNIFK\GTHFSLVDNVKK
						TVSTWLH/SQNAQFYKDRLNGWYHCLQKCL
						QHY*AYVEK
373	1723	A	3181	410	14101	RREVAGPEGKGLLLASAHTMLTPPLLLLLPLL
1						SALVAAAIDAPKTCSPKQFACRDQITCISKGW
						RCDGERDCPDGSDEAPEICPQSKAQRCQPNE
1						HNCLGTELCVPMSRLCNGVQDCMDGSDEGP
						HCRELQGNCSRLGCQHHCVPTLDGPTCYCNS
. !						SFQLQADGKTCKDFDECSVYGTCSQLCTNTD
						GSFICGCVEGYLLQPDNRSCKAKNEPVDRPP
[VLLIANSQNILATYLSGAQVSTITPTSTRQTTA
						MDFSYANETVCWVHVGDSAAQTQLKCARM
			İ		i	PGLKGFVDEHTINISLSLHHVEQMAIDWLTGN
		-				FYFVDDIDDRIFVCNRNGDTCVTLLDLELYNP
]					İ	KGIALDPAMGKVFFTDYGQIPKVERCDMDG
					ĺ	QNRTKLVDSKIVFPHGITLDLVSRLVYWADA
1					l	YLDYIEVVDYEGKGRQTIIQGILIEHLYGLTVF
]				İ		ENYLYATNSDNANAQQKTSVIRVNRFNSTEY
					l	QVVTRVDKGGALHIYHQRRQPRVRSHACEN
						DQYGKPGGCSDICLLANSHKARTCRCRSGFS
					ļ	LGSDGKSCKKPEHELFLVYGKGRPGIIRGMD MGAKVPDEHMIPIENLMNPRALDFHAETGFI
				1	l	YFADTTSYLIGROKIDGTERETILKDGIHNVE
					l	GVAVDWMGDNLYWTDDGPKKTISVARLEK
			! I	ľ	ł	AAQTRKTLIEGKMTHPRAIVVDPLNGWMYW
		1		ŀ		TDWEEDPKDSRRGRLERAWMDGSHRDIFVT
						SKTVLWPNGLSLDIPAGRLYWVDAFYDRIETI
[ĺ	LLNGTDRKIVYEGPELNHAFGLCHHGNYLFW
			<u> </u>			TEYRSGSVYRLERGVGGAPPTVTLLRSE\RPPI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uchee		914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline,
Dence	f	ĺ	714	_		Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	i :			peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
		1				FEIR\MYDAQHQQVGSNKCRVNNAGCSSLCL
1 :	!	ł				ATPGSRQCACAEDQVLDADGVTCLANPSYVP
						PPQCQPGEFACANSRCIQERWKCDGDNDCLD
1 .						NSDEAPALCHQHTCPSDRFKCENNRCIPNRW
	! .					LCDGDNDCGNSEDESNATCSARTCPPNQFSC
	l i					ASGRCIPISWTCDLDDDCGDRSDESASCAYPT
						CFPLTQFTCNNGRCININWRCDNDNDCGDNS
	li					DEAGCSHSCSSTQFKCNSGRCIPEHWTCDGD
)	1					NDCGDYSDETHANCTNQATRPPGGCHTDEF
	i l		i			QCRLDGLCIPLRWRCDGDTDCMDSSDEKSCE
	!					CALTAICOBEANCECCADE YBCICA YARRODCD
1						GVTHVCDPSVKFGCKDSARCISKAWVCDGD
			-	l		NDCEDNSDEENCESLACRPSHPCANNTSVC
1				ļ		LPPDKLCDGNDDCGDGSDEGELCDQCSLNN
				j		GGCSHNCSVAPGEGIVCSCPLGMELGPDNHT
				ļ		CQIQSYCAKHLKCSQKCDQNKFSVKCSCYEG
Į į				3		WVLEPDGESCRSLDPFKPFIIFSNRHEIRRIDLH
	i					KGDYSVLVPGLRNTIALDFHLSQSALYWTDV
1		·		i		VEDKIYRGKLLDNGALTSFEVVIQYGLATPEG
			}			LAVDWIAGNIYWVESNLDQIEVAKLDGTLRT
1		1				TLLAGDIEHPRAIALDPRDGILFWTDWDASLP
				•		RIEAASMSGAGRRTVHRETGSGGWPNGLTV
]		١		J		DYLEKRILWIDARSDAIYSARYDGSGHMEVL
	1					RGHEFLSHPFAVTLYGGEVYWTDWRTNTLA
			i			KANKWTGHNVTVVQRTNTQPFDLQVYHPSR
))	l J					QPMAPNPCEANGGQGPCSHLCLINYNRTVSC
•						ACPHLMKLHKDNTTCYEFKKFLLYARQMEIR
						GVDLDAPYYNYIISFTVPDIDNVTVLDYDARE
		j)		QRVYWSDVRTQAIKRAFINGTGVETVVSADI.
1	i	ĺ	}	.	.	PNAHGLAVDWVSRNLFWTSYDTNKKQINVA
			1			RLDGSFKNAVVQGLEQPHGLVVHPLRGKLY
]		J	}	ļ		WTDGDNISMANMDGSNRTLLFSGQKGPVGL
1 1			1			AIDFPESKLYWISSGNHTINRCNLDGSGLEVID
			İ	1		AMRSQLGKATALAIMGDKLWWADQVSEKM
, ,	j		ļ	j		GTCSKADGSGSVVLRNSTTLVMHMKVYDESI
!			1	i		QLDHKGTNPCSVNNGDCSQLCLPTSETTRSC
i l			[MCTAGYSLRSGQQACEGVGSFLLYSVHEGIR
] [l	j	l		GIPLDPNDKSDALVPVSGTSLAVGIDFHAEND
ļ·	ĺ			[[TIYWVDMGLSTISRAKRDQTWREDVVTNGIG
1	.			ŀ		RVEGIAVDWIAGNIYWTDQGFDVIEVARLNG
}		J		l		SFRYVVISQGLDKPRAITVHPEKGYLFWTEW
ļ [ĺ			Į	[GOYPRIERSRLDGTERVVLVNVSISWPNGISV
}	l	1		j	ļ	DYODGKLYWCDARTDKIERIDLETGENREVV
]			- 1	ĺ		LSSNNMDMFSVSVFEDFIYWSDRTHANGSIK
		ļ	- 1	ſ	ĺ	RGSKDNATDSVPLRTGIGVQLKDIKVFNRDR
		ŀ	1			
	- 1	.		ļ		QKGTNVCAVANGGCQQLCLYRGRGQRACA
{	- (ĺ	- 1	l	{	CAHGMLAEDGASCREYAGYLLYSERTILKSI
l í	ļ	- 1	-	ļ	į	HLSDERNLNAPVQPFEDPEHMKNVIALAFDY
		1	}	1	ļ	RAGTSPGTPNRIFFSDIHFGNIQQINDDGSRRIT
	ſ	1	1	ľ	Ī	IVENVGSVEGLAYHRGWDTLYWTSYTTSTIT
		ŀ	- 1	j	l	RHTVDQTRPGAFERETVITMSGDDHPRAFVL
]]	}	i		l	DECQNLMFWTNWNEQHPSIMRAALSGANVL
		1	1		İ	TLIEKDIRTPNGLAIDHRAEKLYFSDATLDKIE
			ļ		1	RCEYDGSHRYVILKSEPVHPFGLAVYGEHIF
			ļ		ł	WTDWVRRAVQRANKHVGSNMKLLRVDIPQ
			1	1		QPMGIIAVANDTNSCELSPCRINNGGCQDLCL
•			•		ł	LTHQGHVNCSCRGGRILQDDLTCRAVNSSCR
		.	1		Ì	AQDEFECANGECINFSLTCDGVPHCKDKSDE
ĺĺ	- 1	[1		ı	KPSYCNSRRCKKTFRQCSNGRCVSNMLWCN
]	ļ			ļ	GADDCGDGSDEIPCNKTACGVGEFRCRDGTC
L					i	IGNSSRCNQFVDCEDASDEMNCSATDCSSYF

000.00	ano m	1 84	T-6700	- ·		
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	лоа	in NO:	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-		USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline.
uence	donot		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
				1		RLGVKGVLFQPCERTSLCYAPSWVCDGAND
			1			CGDYSDERDCPGVKRPRCPLNYFACPSGRCIP
						MSWTCDKEDDCEHGEDETHCNKFCSEAQFE
						CONHRCISKOWLCDGSDDCGDGSDEAAHCE
1						GKTCGPSSFSCPGTHVCVPERWLCDGDKDCA
						DGADESIAAGCLYNSTCDDREFMCQNRQCIP
						KHFVCDHDRDCADGSDESPECEYPTCGPSEF
				1		RCANGRCLSSRQWECDGENDCHDQSDEAPK
						NPHCTSPEHKCNASSQFLCSSGRCVAEALLCN
						GQDDCGDSSDERGCHINECLSRKLSGCSQDC
						EDLKIGFKCRCRPGFRLKDDGRTCADVDECS
						TTFPCSQRCINTHGSYKCLCVEGYAPRGGDP
						HSCKAVTDEEPFLIFANRYYLRKLNLDGSNY
						TLLKQGLNNAVALDFDYREQMIYWTDVTTQ
						GSMIRRMHLNGSNVQVLHRTGLSNPDGLAV
			i			DWVGGNLYWCDKGRDTIEVSKLNGAYRTVL
						VSSGLREPRALVVDVQNGYLYWTDWGDHSL
						IGRIGMDGSSRSVIVDTKITWPNGLTLDYVTE
						RIYWADAREDYIEFASLDGSNRHVVLSQDIPH
					:	IFALTLFEDYVYWTDWETKSINRAHKTTGTN
						KTLLISTLHRPMDLHVFHALRQPDVPNHPCK
						VNNGGCSNLCLLSPGGGHKCACPTNFYLGSD
						GRTCVSNCTASQFVCKNDKCIPFWWKCDTE
				. 1		DDCGDHSDEPPDCPEFKCRPGQFQCSTGICTN
						PAFICDGDNDCQDNSDEANCDIHVCLPSQFK CTNTNRCIPGIFRCNGQDNCGDGEDERDCPE
						VTCAPNQFQCSITKRCIPRVWVCDRDNDCVD
						GSDEPANCTQMTCGVDEFRCKDSGRCIPARW
				•		KCDGEDDCGDGSDEPKEECDERTCEPYQFRC
				,		KNNRCVPGRWQCDYDNDCGDNSDEESCTPR
			' I		ĺ	PCSESEFSCANGRCIAGRWKCDGDHDCADGS
			ĺ			DEKDCTPRCDMDQFQCKSGHCIPLRWRCDA
						DADCMDGSDEEACGTGVRTCPLDEFQCNNT
						LCKPLAWKCDGEDDCGDNSDENPEECARFV
						CPPNRPFRCKNDRVCLWIGRQCDGTDNCGD
		[1			GTDEEDCEPPTAHTTHCKDKKEFLCRNQRCL
j						SSSLRCNMFDDCGDGSDEEDCSIDPKLTSCAT
	1		1	1		NASICGDEARCVRTEKAAYCACRSGFHTVPG
	1		i	İ		QPGCQDINECLRFGTCSQLCNNTKGGHLCSC
	ļ			-		ARNFMKTHNTCKAEGSEYQVLYIADDNEIRS
	i				1	LFPGHPHSAYEQAFQGDESVRIDAMDVHVKA
ł	İ	i	1	İ	ł	GRVYWTNWHTGTISYRSLPPAAPPTTSNRHR
				İ		RQIDRGVTHLNISGLKMPRGIAIDWVAGNVY
- 1		- 1		l		WTDSGRDVIEVAQMKGENRKTLISGMIDEPH
ţ	1	Ì		l		AIVVDPLRGTMYWSDWGNHPKIETAAMDGT
İ	1	ł		ĺ		LRETLYQDNIQWPTGLAVDYHNERLYWADA
ĺ	ĺ	ſ	ļ	1	ľ	KLSVIGSIRLNGTDPIVAADSKRGLSHPFSIDV
	1	1	I	ļ		FEDYIYGVTYINNRVFKIHKFGHSPLVNLTGG
ļ	1	1	Ī			LSHASDVVLYHQHKQPEVTNPCDRKKCEWL
		ĺ		1	1	CLLSPSGPVCTCPNGKRLDNGTCVPVPSPTPP
		ļ	l	i		PDAPRPGTCNLQCFNGGSCFLNARRQPKCRC
[ſ	•	[1	QPRYTGDKCELDQCWEHCRNGGTCAASPSG
	ŀ	1	}		l	MPTCRCPTGFTGPKCTQQVCAGYCANNSTCT
	ļ	Ī	İ	ł	ļ	VNQGNQPQCRCLPGFLGDRCQYRQCSGYCE
	1			ŀ		NFGTCQMAADGSRQCRCTAYFEGSRCEVNK
	l		l	ł		CSRCLEGACVVNKQSGDVTCNCTDGRVAPS
1	[[[[ĺ	CLTCVGHCSNGGSCTMNSKMMPECQCPPHM TGPRCEEHVFSQQQPGHIASILIPLLLLLLLVL
1		- 1	- 1			TOT MODERN ALPÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄÄ
	ł	- 1	- 1	I		VAGVVFWVKRRVOGAYGEGUGDAMAG
.			- 1			VAGVVFWYKRRVQGAKGFQHQRMTNGAM NVEIGNPTYKMYEGGEPDDVGGLLDADFAL

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO:	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
į	ļ	}	ļ	peptide		/=possible nucleotide deletion, \=possible
		<u> </u>		sequence		nucleotide insertion
						DPDKPTNFTNPVYATLYMGGHGSRHSLASTD EKRELLGRGPEDEIGDPLA
374	1724	A	3187	191	1815	CLELASAGKIPEESKALSLLAPAPTMTSLMPG
3/4	1724	^	3107	131	10.5	AGLLPIPTPNPLTTLGVSLSSLGAIPAAALDPNI
						ATLGEIPQPPLMGNVDPSKIDEIRRTVYVGNL
1		1	ļ	}		NSQTTTADQLLEFFKQVGEVKFVRMAGDET
			ļ	1		QPTRFAFVEFADQNSVPRALAFNGVMFGDRP LKINHSNNAIVKPPEMTPQAAAKELEEVMKR
						VREAOSFISAAIEPGWLHSTSLCNDFLGCF*RR
ł			ļ	ł	1	RMYRE*APCTICGTFHLCLINWDL*LF*AYTA
		1		Ì		K*FFPPRVWKEQ*KKRR\RSRSHTRSKSRSSSK
			ļ	1	ļ	SHSRRKRSQSKHRSRSHNRSRSRQKDRRRSK SPHKKRSKSRERRKSRSRSHSRDKRKDTREKI
1	Į	l	{	ł		KEKERVKEKDREKEREKEREKEKERGKN
				İ		KDRDKEREKDREKDKEKDREREREKEHEKD
	· ·	1	ļ			RDKEKEKEQDKEKEREKDRSKEIDEKRKKDK
ł	1	ĺ		İ	i	KSRTPPRSYNASRRSRSSSRERRRRSRSSSRS
1	8			1		PRTSKTIKRKSSRSPSPRSRNKKDKKREKERD HISERRERERSTSMRKSSNDRDGKEKLEKNST
'						S
375	1725	A	3192	415	101	AHSSHQTRAILQEFQWDIIRHPPL\SPNLALSG
	ł			l	1	F\FPNLKKSLRGTHFSSVKK\TTLTWLNSQDP
		i		1		WF/FFYP*SPDLQIPSSFRNGLNDWYHHSQKC PDLDGAYVKK
376	1726	A	3199	931	418	GV*WCDLGSPQPPPPGFKQFCLGRSSSWDYR
370	1720	 ^	3177	/31		HVPPHPANFVFLLETGFLHAGQAGL\GDPPAS
			1	'	1	ASQSAGITGVSHTWPKNHLIFYACLVIRSKRI
	<u> </u>		1	1	1206	K KTGYTSRGSPLSPQSSIDSELSTSELEDDSISM
377	1727	A	3201	274	1285	GYKLQDLTDVQIMARLQEESLRQDYASTSAS
		1				VSRHSSSVSLSSGKKGTCSDQEYDQYSLEDEE
						EFDHLPPPQPRLPRCSPFQRGIPHSQTFSSIREC
1		1	[RRSPSSQYFPSNNYQQQQYYSPQAQTPDQQP
				Ì		NRTNGDK/PPKKYA*PSPDAKYNCH**QH\SSP VTVRNSQSFDSSLHGAGNGISRIQSCIPSPGQL
1]		j	1	QHRVHSVGHFPVSIRQPLKATAYVSPTVQGSS
						NMPLSNGLQLYSNTGIPTPNKAAASGIMGRS
	1					ALPRPSLAINGSNLPRSKIAQPVRSFLQPPKPL
270	1720	 	3202	112	1789	SSLSTLRDGNWRDGCY VPGVTESRPSVLRGDHLFALLSSETHQEDPIT
378	1728	A	3202	112	1705	YKGFVHKVELDRVKLSFSMSLLSRFVGWG*
			1			PFKVNFY/TFNRQPLRV\QHRALELTGRWLLW
	1			1		PMLFP\VAPRDVPLLPSDVKLKLYDRSLESNP
	j	}		ļ		EQLQAMRHIVTGTTRPAPYIIFGPPGTGKTVT LVEAIKQVVKHLPKAHILACAPSNSGADLLC
						QRLRVHLPSSIYRLLAPSRDIRMVPEDIKPCCN
				1		WDAKKGEYVFPAKKKLQEYRVLITTLITAGR
	1	1	1			LVSAQFPIDHFTHIFIDEAGHCMEPESLVAIAG
			1	1		LMEVKETGDPGGQLVLAGDPRQLGPVLRSPL TQKHGLGYSLLERLLTYNSLYKKGPDGYDPQ
		1				FITKLLRNYRSHPTILDIPNQLYYEGELQACA
	1	1				DVVDRERFCRWAG\LPRQGFPIIFHGVMGKD
	1		1			EREGNSPSFFNPEEAATVTSYLKLLLAPSSKK
		1				GKARLSPRSVGVISPYRKQVEKIRYCITKLDR ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP
}	1		1	1		ELRGLDDIKDLKVTCCSTVTPCLPCAPTCPLP ETSSSFHSSPRPRPTPAALNRARALPEPLTPGD
	1				1	SNLRVWDGIRKPACLTNTSCHS
379	1729	A	3206	432	130	PKAAPSVXLWFPPFL*GSFKPTKGHTXCVXIK
L	<u></u>	1		<u> </u>	l	*LSTREAXDSXPGRQIAXXRQGGKVETTTAL

SEQ ID NO: of nucl- nucl- eotide seq- uence SEQ ID NO: of nucl- eotide seq- uence SEQ ID NO: of nucl- seq- uence SEQ ID NO: of nucl- seq- uence SEQ ID NO: of nucl- seq- uence SEQ ID NO: of nucl- seq- uence SEQ ID NO: of nucl- beginning nucleotide location corresponding to last amino acid residue of peptide sequence SEQ ID NO: of nucl- seq- uence SEQ ID NO: of nucl- seq- uence SEQ ID NO: of nucl- beginning nucleotide location corresponding to last amino acid residue of peptide sequence SEQ ID NO: of nucl- seq- uence SEQ ID NO: of nucl- seq- uence SEQ ID NO: of nucl- seq- seq- uence SEQ ID NO: of nucleotide location seq- seq- uence SEQ ID NO: of nucleotide location seq- seq- uence SEQ ID NO: of nucleotide location seq- seq- seq- uence SEQ ID NO: of nucleotide location seq- seq- seq- uence SEQ ID NO: of nucleotide location seq- seq- seq- uence SEQ ID NO: of nucleotide location seq- seq- seq- seq- uence SEQ ID NO: of nucleotide location seq- seq- seq- seq- seq- seq- seq- seq-	dine, line, an, lon, ble GSHVQYRC WSDRVPKC HYQAGESL WTSQPPLC LGLVIVLG
nucleotide sequence Deptide sequence Deptide s	an, lon, ble GSHVQYRC WSDRVPKC HYQAGESL WTSQPPLC LGLVIVLG
eotide sequence USSN 09/496 corresponding uence uence USSN 09/496 of peptide sequence USSN 09/496 of peptide sequence USSN 09/496 of peptide sequence USSN 09/496 of peptide sequence USSN 09/496 of peptide sequence USSN 09/496 of peptide sequence IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	an, lon, ble GSHVQYRC WSDRVPKC HYQAGESL WTSQPPLC LGLVIVLG
sequence Sequence Uence 09/496 corresponding to last amino acid residue of peptide sequence 09/496 of peptide s	an, lon, ble GSHVQYRC WSDRVPKC HYQAGESL WTSQPPLC LGLVIVLG
uence 914 ng to first amino acid residue of peptide sequence peptide sequence 914 ng to first amino acid residue of peptide sequence peptide sequence 914 ng to first amino acid residue of peptide sequence peptide sequence 914 ng to first amino acid residue of peptide sequence peptide sequence 914 ng to first amino acid residue of peptide sequence 92Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptoph Y=Tyrosine, X=Unknown, *=Stop corposible nucleotide deletion, \=possi nucleotide insertion MTCADPGEIANGHRTASDAGFPV LPGYSLEGAAMLTCYSRDTGTPK: ALKYEPCLNPGVPENGYQTLYKH RFFCYEGFELIGEVTITCVPGHPSQ KVTQTTDPSRQLEGGNLALAILLP	an, lon, ble GSHVQYRC WSDRVPKC HYQAGESL WTSQPPLC LGLVIVLG
amino acid residue of peptide sequence of peptide sequence	lon, ble GSHVQYRC WSDRVPKC HYQAGESL WTSQPPLC LGLVIVLG
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop con /=possible nucleotide deletion, \=possi nucleotide insertion MTCADPGEIANGHRTASDAGFPV. LPGYSLEGAAMLTCYSRDTGTPK: ALKYEPCLNPGVPENGYQTLYKH RFFCYEGFELIGEVTITCVPGHPSQ KVTQTTDPSRQLEGGNLALAILLP	lon, ble GSHVQYRC WSDRVPKC HYQAGESL WTSQPPLC LGLVIVLG
peptide /=possible nucleotide deletion, \=possi sequence nucleotide insertion MTCADPGEIANGHRTASDAGFPV LPGYSLEGAAMLTCYSRDTGTPK ALKYEPCLNPGVPENGYQTLYKH RFFCYEGFELIGEVTITCVPGHPSQ KVTQTTDPSRQLEGGNLALAILLP	GSHVQYRC WSDRVPKC HYQAGESL WTSQPPLC LGLVIVLG
sequence nucleotide insertion MTCADPGEIANGHRTASDAGFPV LPGYSLEGAAMLTCYSRDTGTPK ALKYEPCLNPGVPENGYQTLYKH RFFCYEGFELIGEVTITCVPGHPSQ KVTQTTDPSRQLEGGNLALAILLP	GSHVQYRC WSDRVPKC HYQAGESL WTSQPPLC LGLVIVLG
MTCADPGEIANGHRTASDAGFPV LPGYSLEGAAMLTCYSRDTGTPK ALKYEPCLNPGVPENGYQTLYKH RFFCYEGFELIGEVTITCVPGHPSQ KVTQTTDPSRQLEGGNLALAILLP	WSDRVPKC HYQAGESL WTSQPPLC LGLVIVLG
LPGYSLEGAAMLTCYSRDTGTPK: ALKYEPCLNPGVPENGYQTLYKH RFFCYEGFELIGEVTITCVPGHPSQ KVTQTTDPSRQLEGGNLALAILLP	WSDRVPKC HYQAGESL WTSQPPLC LGLVIVLG
ALKYEPCLNPGVPENGYQTLYKH RFFCYEGFELIGEVTITCVPGHPSQ KVTQTTDPSRQLEGGNLALAILLP	HYQAGESL WTSQPPLC LGLVIVLG
RFFCYEGFELIGEVTITCVPGHPSQ KVTQTTDPSRQLEGGNLALAILLP	WTSQPPLC LGLVIVLG
KVTQTTDPSRQLEGGNLALAILLP	LGLVÌVLG
KVTQTTDPSRQLEGGNLALAILLP SGVYIYYTKI OGKSI FGFSGSHSV	LGLVIVLG
SGVYIYYTKI.OGKSI FGFSGSHSV	CDYM MOSA
The state of second in the second in t	SPITVESDF
SNPLYEAGDTREYEVSI	
386 1736 A 3250 5725 3984 GTSTVTMATKKHFSILNLLGMLL	KKDNQDT
RKLLMTWALEVAVVMKKSETYA	PLFCLPSF
HKFCKGLLADTLVEDVNICLQACS	SSLHALSSS
LPDDLLQRCVDVCRVQLVHRGTC	IROAFGKL
LKSIPLGVFLSNNNHTEIQEISLALI	RSHMSKAP
SNTFHPQDFSD/VISFILYGNSHRTC	KDNWLE
RLFYSCQRLDKRDQSTIPRNLLKTI	DAVLWOW
AIWEAAQFTVLSKLRTPLGRAQDT	FOTIEGIR
SLAGHTLNPDQDVSQWTTADNDE	GHGNNOL
RLVLLLQYLENLEKLMYNAYEGC	ANALTSPP
KVIRTFLYTNRQTCQDWLTRIRLS	MRVGLLA
GQPAVTVRHGFDLLTEMKITSLSC	GNELEVSI
MMVVEALCELHCPEAIQGIAVWS	SIVGKHI
LWINSVAQQAEGRFEKASVEYQEI	I CAMTG
VDCCISSFDKSVLTLASAGCKSASI	KHCI NGE
SRKSVLSKPTDSSPEVINYLGNKAC	TECVICTA
DWAAVQEWQNAIHDLKKSTSSTS	DECISIA
NYIKSLSSFESGKFVECTEQLELLPO	CIVERADI
GGSKEKIDMKKLLRNM	JENINLLA
387 1737 A 3255 380 76 MDIFLYNCKYQVQTEI*NSIQHIMA	VENT CDE
LKYVHNL*AENYKTLMK*INEDLN	
S*TARLNKMSIPTKTIFRFKAIYIKII	KQKDVPY
NMO	AITFIEI
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	STILAQIG
FELLDSSDLPASASKSAGITCMSHH	ARTLSLK
389 1739 A 3269 1 332 LDGYHTPIYMI NRIIRI PA AL *IISD	
LEGITITIVILIKIIKLIKALIISE	QTGHALTI
LTRLETQMINADYQNKLTLDYLLT	TDREVYE
PFNLTNYCLHIHNQRLGAYDLG*V	*Q/KLAHV
390 1740 A 3270 2 372 GPCHDONK GKSIDGPDAOAFACGG	
372 OKCIDQINGKSIDGI DAQAEACGO	
LVNQNPIGQPLACRRLTRKIYEGIK	KAVKPNH
SPRGVKKVHKFVNKGEKGIMVLAG	GDTLGIGV
YCLLPCMC*DRKLTYAHIPSTTDLC	GAGAGY
391 1741 A 3273 1 187 FFQEMLDIMKAISDMMGKCTYPVI	KEDAPRO
HVETFFQ\EELTRSQEGMKLGENFL	MFAMPP
DDSKESKGK*FFQEMLDIMKAISDM	MGKCTY
PVLKEDAPROHVETFFOVGINOKSI	GHEVRR
KFPDVCHAPR KFPDVCHAPR	
392 1742 A 3281 901 521 FFFGDGVSPCRQAGV*WHDLDSLQ	NLPPGFK
RFSYLSLPSSWIDYRHVLPRQANFO	
FTMLARMVSIS*PRDLPALASQSAG	
APPQMDFTFALLCFALKGCLPRQK	EGGTI NI I
393 1743 A 3283 385 3 RNRSVVPEFVLLGLSAGPQTQTLLF	
LLTVMGNLLLLVVINADSCLHTPM	
SFLDLCHSSVTAPKLLENLLSEKKT	
A*VFFVFATGGTESSLLAVMAYDR	IVAIKIK
204	VI III A
CIRCRADEDICTINATICIACASOF	
LDNCPEGLEANNHTMECVSIVHCE	
WSPCTKKGKTCGFKRGTETRVREII	
NLCPPTNETRKCTVQRKKCQKGER	UKKGRE

CEO ID	SEQ ID	1 1/24	Leco	D. J. J.	1 5 1 1 1	
SEQ ID	NO: of	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	peptide	hod	ID NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-	l	USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
401100	1		 '''	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1		ł	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	j	1	ļ	peptide	,	/-possible nucleotide deletion, \-possible
				sequence		nucleotide insertion
	<u> </u>	T	i .			RKRKKPNKGESKEAIPDSKSLESSKEIPEQREN
			1	Ì		K000
395	1745	A	3286	1	340	RVLYVPSMGFCILVAHGWQKISTKSVFKKLS
		l		_		WICLSMVILTHSLKTFHRNWDWESEYTLFMS
						ALKVNKNNAKLWNNVGHALENEKNFERAL
	ļ	ļ				KYFLQATHVQPDDIGAHMNVGR
396	1746	Α	3293	1	172	GFRAVVMTVKTEAAKGTLTYSRMRGMVAIL
İ					}	IAFMKQRRMGLNDFIQKIANNSYACKQ
397	1747	Α	3295	12	401	AEPACGASSCTPPSLRSSSSQSVGPLRPGRPL
		ĺ	j			WSEACAFL*AAAPQGPASPCCGLPSGFPRVW
		ļ			Ì	AQCCPPGGALRFPEGLGSVLSPRRCPQVSRGS
		1			ł	GLSAVPQEVPSGFLGPGLRACPQEAPSRFLRA
					ŀ	GLT
398	1748	A	3300	1912	2768	KQRRWQNIQRKGPKRYIVIAGNSQSHQPMIFS
			Į į			MLRKLPKVTCRDVLPEIRAICIEEIGCWMQSY
	Į.				· ·	STSFLTDSYLKYIGWTLHDKHREVRVKCVKA
						LKGLYGNRDLTARLELFTGRFKDWMVSMIV
	1	l	l i			DREYSVAVEAVRLLILILKNMEGVLMDVDCE
		ļ				SVYPIV*ASN*GLASAVGEFLYWKLFYPECEI
i	1	İ				RTMGGREQRQSPGAQRTFFQLLLSFFVESKSH
		Ì	!			SVTQAGVQWQFSAHRDLCLPGSSNSHVSASR
İ	İ	(VAGIAGAHRHTWLIYVFFSWRQGFAVLAGL
		i]			VSNS
399	1749	Α	3301	536	2391	LRSYGCKAPSRISHLHK\FLFLLLPSLLMGYSE
)			SPPPITDSWAPFISLTHHVLSQSQSPLSSNCWI
1	}					CLSTHTQ*FTALPADLLTWTQSNVSLHISYLAI
İ						PFLADSFLKPV/L*PGNSAKHLSFKLSSLSMVS
						GRAVALLHLIASGLTSIQTNTASSKPPIWGY\L
	1			,		STQTSFISPPPLCLSRTYPNPAHATMVGQVPQ
			1			SLCGLIFTL/RTPCRPSILHPNYKIISTSAWQKV
						LCFSGSPTIHTSLHLTTGSSFLSFHPIPGFPAAN
						SALYVSSLKGPPGKNVTIPSPVTGT*QPPHRGS
						N/RLTVDKDNFFLSPKPNSLHQLPSQ\TPYQAL
					1	TGAALAGSYPIWENENTLSWLPTFTYNFCLST
						PSLFFLCDTN*YLCLPANWSGTCTLVFQAPTI
	Ì		1			NILPPNQTILISVEASISSSPIRNKWALHLITLLT
						GLGITAALGTGIAGITTSITSYQTLFTTLSNTVE
						DMHTSITSLQRQLDFLVGVILQNWRVLDLLT
						TEKGGTCIYLQEECCFCVNESGIVHIAVRRLH
						DRAAEL*HQVADSWWQGSSLLRWIPWVAPF
1						LGPLIFLFLLLMIGPCIFNLVSRFISQRLNCFIQ
						ASMQKHIDNIFHLCHV*YQSLRGNHSEAPEPR
400	1750				450	P
400	1750	A	3303_	2	453	THWRHSSGVPGSTTARRRRELEIATSDNQE
						YYNRLCQEVTNRERNDQKMLADLDDLNRTK
	, ,				İ	KYLEERLIELLRDKDALWQKSDALEFQQKLS
[AEERWLGDTEANHCLDCKREFSWMVRRHHC
L	1000	_	1			RICGRIFCYYCCNNYVLSKHGGKKERCC
401	1751	A	3304	1	626	MAPQHSSLDDKVPQQASTVCFEFQDILQHSQ
1						CTEHKDSLWGPGARSQPFGAHNTRLSPDSCP
						EKIVLRALKDSRAGMPEQDKDPGVQENPDD
1					l	QRRVPQGTGDAPSAFRPLWDNGGLSPFVSRP
				į		GPLERDLHAQRSEVTYNQRSQSSWMSSFPKR
				ļ		NAFVSPYSSMGQAQP/GLPKTNPIGESCCWEG
لــــا						LSLSTQILG*QKPSKYIPSLCKR
402	1752	Α	3305	1678	172	MELPSGPGPERLFDSHRLPGDCFLLLVLLLYA
				1		PVGFCLLVLRLFLGIHVFLVSCALPDSVLRRF
						VVRTMCAVLGLVARQEDSGLRDHSVRVLISN
						HVTPFDHNIVNLLTTCSTVSESEAESATGRFP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide scq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLLCTPVGL\SRMFTVMGQLLVKPTILEDLDE
						QIYIITLEEALQRPTKWAVFIRW/KYNIMELE QELENVKTLKTKLERRKKASAWERNLVYPA VMVLLLIETSISVLLVACNILCLLVDETAMPK GTRGPGIGNASLSTFGFVGAALEIILIFYLMVS SVVGFYSLRFFGNFTPKKDDTTMTKIIGNCVS ILVLSSALPVMSRTLGITRFDLLGDFGRFNWL GNFYIVLSYNLLFAIVTTLCLVRKFTSAVREE LFKALGLHKLHLPNTSRDSETAKPSVNGHQK AL
410	1760	A	3339		1433	GSHRFSLASPLDPEVGPYCDTPTMRTLFNLL WLALACSPVHTTLSKSDAKKAASKTLLEKSQ FSDKPVQDRGLVVTDLKAESVVLEHRSYCSA KARDRHFAGDVLGYVTPWNSHGYDVTKVFG SKFTQISPVWLQLKRRGREMFEVTGLHDVDQ GWMRAVRKHAKGLVP*CLGSCLRTGLTMISG/ YVLDSEDEIEELSKTVVQVAKNQHFDGFVVE VWNQLLSQKRVGLIHMLTHLAEALHQARLL ALLVIPPATTPGTDQLGMFTHKEFEQLAPVLD GFSLMTYDYSTAHQPGPNAPLSWVRACVQV LDPKSKWRSKILLGLNFYGMDYATSKDAREP VVGARYIQTLKDHRPRMVWDSQVSEHFFEY KKSRSGRHVVFYPTLKSLQVRLELARELGVG VSIWELGQGLDYFYDLL*VGIAASAVDVFFSK PWSE
411	1761	A	3342	74	2701	VATRKLAKGFTQFAKMTEGTKKTSKKFKFK FKGFGSFSNLPRSFTLRRSSASISRQSHLEPDTF EATQDDMVTVPKSPPAYARSSDMYSHMGTM PRPSIKKAQNSQAARQAQEAGPKPNLVPGGV PDPPGLEAAKEVMVKATGPLEDTPAMEPNPS AVEVDPIRKPEVPTGDVEEERPPRDVHSERAA GEPEAGSDYVKFSKEKYILDSSPEKLHKELEE ELKLSSTDLRSHAWYHGRIPREVSETLVQRN GDFLIRDSLTSLGDYVLTCRWRNQALHFKIN KVVVKAGESYTHIQYLFEQESFDHVPALVRY HVGSRKAVSEQSGAIIYCPVNRTFPLRYLEAS YGLGQGSSKPASPVSPSGPKGSHMKRRSVTM TDGLTADKVTRSDGCPTSTSLPRPRDSIRSCA LSMDQIPDLHSPMSPISESPSSPAYSTVTRVHA APAAPSATALPASPVARRSSEPQLCPGSAPKT HGESDKGPHTSPSHTLGKASPSPSLSSYSDPDS GHYCQLQPPVRGSREWAATETSSQQARSYGE RLKELSENGAPEGDWGKTFTVPIVEVTSSFNP ATFQSLLIPRDNRPLEVGLLRKVKELLAEVDA RTLARHVTKVDCLVARILGVTKEMQTLMGV RWGMELLTLPHGKKLRLDLLERFHTMSIML AVDILGCTGSAEERAALLHKTIQLAAELRGT MGNMFSFAAVMGALDMAQISRLEQTWVTLR QRHTEGAILYEKKLKPFLKSLNEGKEGPPLSN TTFPHVLPLTILLECDSAPPEGPEPWGSTEHGV EVVLAHLEAARTVAHHGGLYHTNAEVKLQG FQARPELLEVFSTEFQMRLLWGSQGASSSQA RRYEKFDKVLTALSHKLEPAVRSSEL
412	1762	A	3347	1	898	IDRAAECRTKPLPMAVSIRGNADSIVACLVLM VLYLIKKRLVACAAVFYGFAVHMKIYPETYI LPITLHLLPDRDNDKSLRQFRYTFQACL*ELL KRLCNRTALMFVAVAGLTFFALSFGFYYEYG WEFLEHTYFYHLTRRDIRHNFSPYFYMLYLT AESKWSFSLGIAAFLPQLILLSAVSFAYYRDL VFCWFLHTSIFVTFNKVCTSQYFLWYLCLLPL

SEO ID	SEQ ID	Met	SEQ	Predicted	Deadisted and	I A-:
NO: of	NO: of	hod	ID NO:	beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in NO.	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-	ĺ	USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	dente		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
aciac.	j	J	/14	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
İ		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i		1		peptide	sequence	/=possible nucleotide deletion, \=possible
				sequence	ł	nucleotide insertion
			 	Sequence	ļ	VMPLVRMPWKRAVVLLMLWFIGQAMWLAP
		1			ĺ	AVVI DEOCEMPTE FINA ACIDEL DICORIO
1		ĺ	ĺ	ĺ	Ì	AYVLEFQGKNTFLFIWLAGLFFLLINCSILIQII
413	1763	A	3361	3	474	SHYKEEPLTERIKYD
413	1703	^	2301	3	474	PIPVRWNSLEGRLLRGYEQHANDGKDYISRN
ŀ						*DLRSWTAADMAAQITKRKWEAEEFAEQIKA
			}			YLEGTCVER/LRTHLENGKETLQLTEQSSQPTI
1	.	l	1	}		PIVGIVAGLVLLGAVVTGAVVSAVMCRKKNS
1	ĺ					GHFLPTDRVSYSEAASSDHAQGSDVSLTACK
414	1774	ļ.,	22.60	1400		V
414	1764	Α	3363	1488	453	HQILELKKKILKTYNPDYDEDLVQEASSEDVL
						GVHMVDKDTERDIEMKRQLRRLRELHLYST
1						WKKYQEAMKTSLGVPQRERDEGSLGKPLCP
[.	{	1				PEILSETLPGSVKKRVCFPSEDHLEEFIAEHLP
1		l				EASNQSLLTVAHADAGTQTNGDLEDLEEHGP
			1			GQTVSEEATEVHMMEGDPDTLAELLIRDVLQ
ļ		İ				ELSSYNGEEE\DPEEVKTSLGVPQRGDLEDLE
						EHVPGQTVSEEATGVHMMQVDPATLAKSDL
ł		l	1 1			EDLEEHVPEQTVSEEATGVHMMQVDPATLA
		l				KQLEDSTITGSHQQMSASPSSAPAEEATEKTK
!			1 1			VEEEVKTRKPKKKTRKPSKKSRWNVLKCWD
						IFNIF
415	1765	Α	3369	431	315	IPWSWVGRLSVRKMSILF*LTYNYNAILNKTP
						PSFSPSL
416	1766	A	3373	42	651	RQEKMGLGEIGASGVLRSMLKERKKQNMKG
				•		NGNVTLTPLLPAVQCGCHLQPAGRSPLPSSHS
						APGLCSPLHPLQPQQEASTCPSGTLQGREKAA
]]]			PGQGRPLCSLWAGGAGA\PGERGAEGRGPSD
						QAPDPKSGPWLFPPGLGAPAEVRLHNVPHNL
1 1						RRPPLP*ARGK*PPNSGCPWSEGRAKQPLSCG
						PKPQCSLPSQVPGDTH
417	1767	A	3382	2	2061	EAQDPRACGPDAGGRFAARDAPGNSLRPPPS
						SPP/GWPGQLRLLPRVPGSELRCGKPERGRLP
[[1			1		ASPPGKIRGWPPGISKRPGLGGRSFPPGFAPRT
						WRPEARGPSVQSLPPIFSPQSAQTTAR*RPGAP
					ì	KNAGRCGGA\RGPRLSLGPPPGPPPAPALPAR
			1			ASAGAGAAAAALAVGGVRGAGGARGTGGY
1						GHCSGR/PTGRTGPGPQGPGPPMPARPR*AS\S
}	'				ļ	TRGSRRGPGSRPARAAAAPRAGDHGRRPVRV
						HLRQHTAV*EPRLGDATAPPGGAAGPGAPAP
				•		R\GPGWDCALLPSPGPRSPRAVGCAEPEIWDP
						SPRRGTSPVPSVRSLRSEPANPRLGLPALLNSY
						PLKGPGLPPPWGPRTQTGHVIITVQPSGSCIEH
}						SKSLD/RGPWGAPPWGPSSSGLCSPKLATAGP
. [1	[İ		PQSWGLCQIGRRRGLGGPGLKRGET/GLL*GC
-		- 1		•		SMDHANRTKGPGVPTSNRCFSHIPG\GDGCSD
]	HSSCEGHPDLHAGREMPAAPGLSELERVRFT
			j	j	1	
	[ĺ	· · · · · · · · · · · · · · · · · · ·	ĺ	1	VGCGGLASGISSASVSGLSPNRAGGPGQGDW
				ŀ	1	EMYPVSWQTQESGGQG/SPKTGR*VGMLQA
			- 1	l		GAGSLQGGTGDGVWGLWEDGP/RG*DSPLPS
		}		j	j	GTGTEP*TPTTSIPFFPQPSGVYPSRATLLPMPS
		I		j	j	Y*ALGPSANKSEKPLLSFLYRGLCCRISLQLA
418	1768	, 	2200	204	3131	KGIGQLSEIPLLNVETAFWSMWVTYFRK
416	1/00	A	3398	304	2121	EEEEEEDEDDDDNNEEEEFECYPPGMKVQV
	1					RYGRGKNQKMYEASIKDSDVEGGEVLYLVH
	I			}		YCGWNVRYDEWIKADKIVRPADKNVPKIKH
	ŀ	J		}	j	RKKIKNKLDKEKDKDEKYSPKNCKPPALGPN
	Į	I		i]	PPFQTNPISWKWYPKLDLTDAKNSDTAHIKSI
		!	- 1	[I	EITSILNGLQASESSAEDSEQEDERGAQDMDN
		i				NGKEESKIDHLTNNRNDLISKEEQNSSSLLEE

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (AmAlonia Co.C.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ucnec	J	914	ng to first	acid residue	
ualce			714	amino acid	of peptide	Q=Glutamine, R=Arginine, S=Serine,
	1	1		residue of	sequence	T=Threonine, V=Valine, W=Tryptophan,
	i	1	<u> </u>	peptide	Sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	İ		i			/=possible nucleotide deletion, \=possible
		<u> </u>		sequence		nucleotide insertion
	Ì				1	NKVHADLVISKPVSKSPERLRKDIEVLSEDTD
İ		}	1		i	YEEDEVTKKRKDVKKDTTDKSSKPQIKRGKR
1	,]	j .		J	RYCNTEECLKTGSPGKKEEKAKNKESLCMEN
1		1		}		SSNSSSDEDEEETKAKMTPTKKYNGLEEKRK
		1		i		SLRTTGFYSGFSEVAEKRIKLLNNSDERLQNS
1						RAKDRKDVWSSIQGQWPKKTLKELFSDSDTE
l		İ				AAASPPHPAPEEGVAEESLQTVAEEESCSPSV
1 .		j .				ELEKPPPVNVDSKPIEEKTVEVNDRKAEFPSS
		i '				GSNFSA*IPLPYLHLNRLHQSL*QKGSRQQSS
		ł				VTVSEPLAPNQEEVRSIKSETDSTIEVDSVAGE
)		ł				LQDLQSERE*LASRF*CQCELKQ**SARTRTS*
<u> </u>						KSLYRSEKSERCSGRRKFIKKAEKKP*SNSGK
						QQKEGK
419	1769	Α	3399	206	463	QRECLSIHIGQAGIQIGDACWELYCLEHGIQP
						NGVVLDTQQDQLENAKMEHTNASFDTFFCE
		ļ i				TRAGKHVPRALFVDLEPTVIDGIR
420	1770	Α	3408	1010	685	RRLSFFF*IWSSVLVTQARVQWRDLGSPQPLP
1 1]				PGFKRFSCLSLPSSWDYRHPSPRPVNF/HVFLV
						VMGFHHVGQAGLELLTSGDLPALASOSARIT
ļ						GVNHCAQPRGHFH
421	1771	Α	3409	355	1326	ADSNLIESCWQELGLGPWGGDWRVEQVGAS
						ASLRFPREVCSIRFLFTAVSLLSLFLSAFWLGL
						LYLVSPLENEPKEMLTLSEYHERVRSOGOOL
						QQLQAELDKLHKEVSTVRAANSERVAKLVF
1						QRLNEDFVRKPDYALSSVGASIDLOKTSHDY
[]						ADRNTAYFWNRFSFWNYARPPTVILEPHVFP
1						GNCWAFEGDQGQVVIQLPGRVQLSDITLOHP
				•		PPSVEHTGGANSAPRDFAVFFLLSFFTHQGLQ
				•		VYDETEVSLGKFTFDVEKSEIQTFHLONDPPA
				'		AFPKVKIQILSNWGHPRFTCLYRVRAHGVRT
						SEGAEGSAQGPH
422	1772	A	3412	2	421	EFDAQPSIGALVVFKRP*ATTGSDPGPKRGMN
			· · · · ·	-	721	YI.VSCSMRSPESGKGEPGTARDYTPMGRPPP
1 1			1			PVPSVSPGPLPGSLAIAPHSPEPHPWEQQPPRG
						QARSPPGGWLGSAT/RVRRPHNHP/RGH/HSP
						VDTAGAPASPGPDVCE
423	1773	Α	3420	91	706	
ا حد	1//3	^	3420]	700	DAQRAIYSSVGPAVSLRQRQQDGAVKESGR/
						RGGVRSFSRAAAAMAPIKVGDAIPAVEVFEG
						EPGNKVNLAELFKGKKGVLFGVPGAFTPGCS
1						KTHLPGFVEQAEALKAKGVQVVACLSVNDA
						FVTGEWGRAHKAEGKVRLLADPTGAFGKET
l i			- 1		1	DLLLDDSLVSIFGNRRLKRFSMVVQDGIVKA
104	1551		2:21			LNVEPDGTGLTCSLAPNIISQL
424	1774	A	3421	4	7688	RQVTRVGTRVLGSTTAAVFLSVEDDNDNAPQ
]]	j		j	į		FSEKRYVVQVREDVTPGAPVLRVTASDRDKG
	-	.	1			SNAVVHYSIMSGNARGQFYLDAQTGALDVV
i I		l l		-		SPLDYETTKEYTLRVRAQDGGRPPLSNVSGL
!						VTVQVLDINDNAPIFVSTPFQATVLESVPLGY
		i				LVLHVQAIDADAGDNARLEYRLAGVGHDFP
}	1	٠	Į.		ļ	FTINNGTGWISVAAELDREEVDFYSFGVEAR
		1		l	J	DHGTPALTASASVSVTALDVNDNNPTFTQPE
				l		YTVRLNEDAAVGTSVVTVSAVDRDAHSVITY
			1	l		QITSGNTRNRFSITSQSGGGLVSLALPLDYKLE
]]				l	ļ	RQYVLAVTASDGTRQDTAQIVVNVTDANTH
		ŀ		l		RPVFQSSHYTVNVNEDRPAGTTVVLISATDE
}	ļ	ļ		l	Ì	DTGENARITYFMEDSIPQFRIDADTGAVTTQA
		1		l		ELDYEDQVSYTLAITARDNGIPQKSDTTYLEI
				l		LVNDVNDNAPQFLRDSYQGSVYEDVPPFTSV
1				l	,	LQISATDRDSGLNGRVFYTFQGGDDGDGDFI
						- (

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						VESTSGIVRTLRRLDRENVAQYVLRAYAVDK GMPPARTPMEVTVTVLDVNDNPPVFEQDEFD VFVEENSPIGLAVARVTATDPDEGTNAQIMY QIVEGNIPEVFQLDIFSGELTALVDLDYEDRPE YVLVIQATSAPLVSRATVHVRLLDRNDNPPV
						LGNFEILFNNYVTNRSSSFPGGAIGRVPAHDP DISDSLTYSFERGNELSLVLLNASTGELKLSR ALDNNRPLEAIMSVLVSDGVHSVTAQCALRV TIITDEMLTHSITLRLEDMSPERFLSPLLGLFIQ AVAATLATPPDHVVVFNVQRDTDAPGGHILN VSLSVGQPPGPGGGPPFLPSEDLQERLYLNRS
						LLTAISAQRVLPFDDNICLREPCENYMRCVSV LRFDSSAPFIASSSVLFRPIHPVGGLRCRCPPGF TGDYCETEVDLCYSRPCGPHGRCRSREGGYT CLCRDGYTGEHCEVSARSGRCTPGVCKNGGT CVNLLVGGFKCDCPSGDFEKPYCQVTTRSFP
						AHSFITFRGLRQRFHFTLALSFATKERDGLLL YNGRFNEKHDFVALEVIQEQVQLTFSAGEST TTVSPFVPGGVSDGQWHTVQLKYYNKPLLG QTGLPQGPSEQKVAVVTVDGCDTGVALRFGS VLGNYSCAA\QGTQGGSKKSLDLTGPLLLGG VPDLPESFPVRMRQFVGCMRNLQVDSRHIDM
				,		ADFIANNGTVPGCPAKKNVCDSKTCHNGGTC VNQWDAFSCECPLGFGGKSCAQEMANPQHF LGSSLVAWHGLSLPISQPWYLSLMFRTRQAD GVLLQAITRGRSTITLQLREGHVMLSVEGTGL QASSLRLEPGRANDGDWHHAQLALGAIGGP
				•		GHAILSFDYGQQRAEGNLGPRLHGLHLSNITV GGIPGPAGGVARGFRGCLQGVRVSDTPEGVN SLDPSHGESINVEQGCSLPDPCDSNPCPANSY CSNDWDSYSCSCDPGYYGDNCTNVCDLNPC EHQSVCTRKPSAPHGYTCECPPNYLGPYCET RIDQPCPRGWWGHPTCGPCNCDVSKGFDPDC
						NKTSGECHCKENHYRPPGSPTCLLCDCYPTG SLSRVCDPEDGQCPCKPGVIGRQCDRCDNPF AEVTTNGCEVNYDSCPRAŒAGIWWPRTRFG LPAAAPCPKGSFGTAVRHCDEHRGWLPPNLF NCTSITFSELKGFAERLQRNESGLDSGRSQQL
						ALLENATOHTAGYFGSDVKVAYQLATRLL AHESTQRGFGLSATQDVHFTENLLRVGSALL DTANKRHWELIQQTEGGTAWLLQHYEAYAS ALAQNMRHTYLSPFTIVTPNIVISVVRLDKGN FAGAKLPRYEALRGEQPPDLETTVILPESVFR ETPPVVRPAGPGEAQEPEELARRQRRHPELSQ
-						GEAVASVIIYRTLAGLLPHNYDPDKRSLRVPK RPIINTPVVSISVHDDEELLPRALDKPVTVQFR LLETEERTKPICVFWNHSILVSGTGGWSARGC EVVFRNESHVSCQCNHMTSFAVLMDVSRRE NGEILPLKTLTYVALGVTLAALLTFFFLTLL
						RILRSNQHGIRRNLTAALGLAQLVFLLGINQA DLPFACTVIAILLHFLYLCTFSWALLEALHLY RALTEVRDVNTGPMRFYYMLGWGVPAFITG LAVGLDPEGYGNPDFCWLSIYDTLIWSFAGP VAFAVSMSVFLYILAARASCAAQRQGFEKKG PVSGLQPSFAVLLLLSATWLLALLSVNSDTILL
. ,						FHYLFATCNCIQGPFIFLSYVVLSKEVRKALK LACSRKPSPDPALTTKSTLTSSYNCPSPYADG RLYQP\YGDSAGSLHSTSRSGKSQPSYIPFLLR EESALNPG\QGPPGLGGIPGR/I.CFLGRFKDQQ H\DS*TRDFDSDLSLEDDQSGSYASTHSSDSEE

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	corresponding to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
			į į	amino acid residue of peptide	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
		ļ		sequence		nucleotide insertion EEEEEEEAAFPGEQGWDSLLGPGAERLPLHS TPKDGGPGPGKAPWPGDFGTTAKESSGNGAP
						EERLRENGDALSREGSLGPLPGSSAQPHKGIL KKKCLPTISEKSSLLRLPLEQCTGSSRGSSASE GSRGGPPSRPPPRQSLQEQLNGVMPIAMSIKA GTVDEDSSGSEFLFFNFLH
425	1775	A	3429	155	1417	GEPAVQSCDCGCTQRSCPWLLVAPGLLSSSSS RAASVREAEDAPLQPASIHPVSQGSRGPEGSL GSAECLPGDPLGARRATRAHSPVPGPPPSLPA AGTAVKRGLQPG*GA/GATSTPGTGAATGGL CGPAWAAPSAVGPCCCCPSISTTPSQMRSARP
						SLGCLPSWAS\PGTEHPPGPQGPGPS*DLCSV* KREFQRGPWAGMVILHRISAADPARAPGPDS NLQSALQQPATGCSEPAAVYSPPIGLWGA**P EYG*PQHSLPG*TAPADR*P\AGIKDRVYSNSI YELLENGQRAGTCVLEYATPLQTLFAMSQYS
						QAGFSREDRLEQAKLFCRTLEDILADAPESQN NCRLIAYQEPADDSSFSLSQEVLRHLRQEEKE EVTVGSLKTSAVPSTSTMSQEPELLISGMEKP LPLRTDFS
426	1776	A	3431	1662	369	AIWWLSWLQHDLLPTPTQVAIDFTASNGDPR SSQSLHCLSPRQPNHYLQALRAVGGICQDYD/ SVGESGAGGNRQGGLAQRIPQLFLLPSDKRFP AFGFGARIPPNFEVG*MRGKEGDGGRVSQAE KAGPHCSRLALTG\SHDFAINFDPENPECEGK
						RGDFHLPRLPADTLHTGAQTPLPRAQLPVPST HPRPVFNEISGVIASYRRCLPQIQLYGPTNVAP INRVAEPAQREQSTGQATKYSVLLVLTDGV VSDMAETRTAIVRASRLPMSIIIVGVGNADFS DMRLLDGDDGPLRCPRGVPAARDIVQFVPFR
						DFKDVSPPGPFRLKDSSASHPPKSDLRLPPFD VLLRTREPSWPP*SPTSPSDDPASPTLPLTPNHI TVPTL\AAPSALAKCVLAEVPRQVVEYYASQ GISPGAPRPCTLATTPSPSP
427	1777	A	3446	. 79	9748	GCQSCWPAWPRLRRRGPASAGARLGRKAPW GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA ASRPEASGDCRAGRETAMATLEKLMKAFESL
						KSFQQQQQQQQQQQQQQQQQQQPPPP PPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPP GPAVAEEPLHRPKKFLSATKKDRVNHCLTIC ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA
	-					ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPQK CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP
				i.	:	KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL
						TLHHTOHODHNVVTGALELLQQLFRTPPPEL LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG
						SAGHDITTEQPRSQHTLQADSVDLASCDLTSS ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD GTDNQYLGLQIGQPQDEDEEATGILPDEASEA
						FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF VLRDEATEPGDQENKPCRIKGDIGQSTDDDS APLYHCVRLLSASFLLTGGKNVLVPDRDVRV SVKALALSCVGAAVALHPESFFSKLYKVPLD

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline
uence	ļ	Ī	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
1]	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
			Ĭ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
				peptide		/=possible nucleotide deletion, \=possible
 			ļ	sequence		nucleotide insertion
1	ľ	ľ		1	i	ADAPAPSSPPTSPVNSRKHRAGVDIHSCSQFL
		1				LELYSRWILPSSSARRTPAILISEVVRSLLVVS
J		Ì		I		DLFTERNQFELMYVTLTELRRVHPSEDEILAQ
İ	ļ		ļ	1	1	YLVPATCKAAAVLGMDKAVAEPVSRLLESTL
	İ	ł	l	1		RSSHLPSRVGALHGILYVLECDLLDDTAKQLI PVISDYLLSNLKGIAHCVNIHSQQHVLVMCAT
		l				AFYLIENYPLDVGPEFSASIIQMCGVMLSGSE
			1			ESTPSIIYHCALRGLERLLLSEQLSRLDAESLV
		į				KLSVDRVNVHSPHRAMAALGLMLTCMYTG
l		i .	l	}		KEKVSPGRTSDPNPAAPDSESVIVAMERVSVL
						FDRIRKGFPCEARVVARILPQFLDDFFPPQDIM
						NKVIGEFLSNQQPYPQFMATVVYKVFQTLHS
			1	ļ		TGQSSMVRDWVMLSLSNFTQRAPVAMATWS
1		l				LSCFFVSASTSPWVAAILPHVISRMGKLEOVD
		1				VNLFCLVATDFYRHQIEEELDRRAFQSVLEV
428	1778	<u> </u>	2442			VAAPGSPYHRLLTCLRNVHKVTTC
420	1//8	A	3449	3	430	NSRPSPSAALVEVLLRSGSTFPHTVSGGWAA
1						WGPWSSCSRDCELGFRVRKRTCTNPEPRNGG
						LPCVGDAAEYQDCNPQACPVRGAWSCWTS
						WSPCSASCGGGHYQRTRSCTSPAPSPGEDICL
429	1779	A	3464	583	3	GLHTEEALCATQACPEGWS
			3401	303	,	DALDRRYLERCHPAAGGWVGEGE*ALCQKT/ RFSGVLEPPLPSLKDGGRFPAWT*RSCSKSLR
					[AAFTSQFFPSRRSRASPGSAP\GNGQNLTEQHP
						CPGSCDPQVLSASWM*VEHRSKFRPPP*NSTI
						PPES/RS*QGGTVQTGQHSSGREAGSWRARGR
]		1				NAGRR*KGGGKIGTKQGAVRARKECRGEMA
			_			SGETDSE
430	1780	A	3473	2802	270	FRMRIFLHCPWNQQMWKIWNLLETSLESCKA
		j				HLSIQKLLKER\Q\QLPVFKHRDSIVETLKRHR
	ļ	į				VVVVAGET\GSGKSTQVPHFLLEDLLLNEWE
i l	ĺ	ľ	1	1	1	ASKCNIVCTQPRRISAVSLANRVCDELGCENG
]		PGGRNSLCGYQIRMESRACESTRLLYCTTGV
		- }	İ			LLRKLQEDGLLSNVS/HMFIVDEV\HER\SVQS
				}		DFLLIILKEILQKRSDLHLILMSATVDSEKFST
	. 1	- 1	1	i	1	YFTHCPILRISGRSYPVEVFHLEDIEETGFVLE
	i	1	i		ŀ	KDSEYCQKFLEEEEEVTINVTSKAGGIKKYQE YIPVQTGAHADLNPFYQKYSSRTQHAILYMN
				-	1	PHKINLDLILELLAYLDKSPQFRNIEGAVLIFL
	.		1			PGLAHIQQLYDLLSNDRRFYSERYKVIALHSI
1	İ	- 1	ľ	-	1	LSTQDQAAAFTLPPPGVRKIVLATNIAETGITI
		- 1	ļ	į		PDVVFVIDTGRTKENKYHESSQMSSLVETFVS
i	j		f	j		KASALQRQGRAGRVRDGFCFRMYTRERFEG
		1		1.		FMDYSVPEILRVPLEELCLHIMKCNLGSPEDF
	-	l	- 1	ſ		LSKALDPPQLQVISNAMNLLRKIGACELNEPK
		- 1	į	ļ		LTPLGQHLAALPVNVKIGKMLIFGAIFGCLDP
	ľ		ł	1		VATLAAVMTEKSPFTTPIGRKDEADLAKSAL
- 1		- 1	1	-		AMADSDHLTIYNAYLGWKKARQEGGYRSEI
1	ł	- 1	1	}		TYCRRNFLNRTSLLTLEDVKQELIKLVKAAGF
İ				- 1		SSSTTSTSWEGNRASQTLSFQEIALLKAVLVA
	ŀ	1		1		GLYDNVGKIIYTKSVDVTEKLACIVETAQGK AQVHPSSVNRDLQTHGWLLYQEKIRYARVY
	1	ļ	·]		LRETTLITPFPVLLFGGDIEVQHRERLLSIDGW
- 1	l	- 1		1		IYFQAPVKIAVIFKQLRVLIDSVLRKKLENPK
			- 1			MSLENDKILQIITELIKTENN
431	1781	A	3474	1	441	FRPAPGHVQP*GGSSAAAGGGLLSHPRPCQQ
	- 1	- 1		1	1	PCPPAPAPSRPRSLGSLGQRVPAALATAAQEL
	ļ	- [J	J		PATLGGDGGKPALTAGEAALPGLHRSGVPAA
	ļ	- 1		ļ	•	AARC*PCT/SRPT*STLSPTQAAWWCRPSRRQ
						QRGEASTGGASGRRCGSCFQV

SEQ ID	SEQ ID	Met	SEQ	Predicted	I Desdies desde	
NO: of	NO: of	hod	ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide	"""	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	i	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
			1	peptide]	/=possible nucleotide deletion, \=possible
420	1.700	<u> </u>	1-2	sequence	<u> </u>	nucleotide insertion
432	1782	Α	3478	416	23	QLRRLTLPNFKTY/YSS*IIEIAWH**KNMQID
		1				QWFRRESPEIDLCKYS*LSFDKEAKAIK/WKE
1		ł	ł	į	Į.	CSLFNKWC/YKNWM/LHVQKKRI*VQTLHPS
	1	1]	QKLK\SKWIKDLNVECRITKLLDQEYPGDLGY
433	1783	A	3504	1876	552	SRALNSGSR
.55	1703	'	3304	1870	332	CLAPCSPQPEKNGMQPLLLLLPPLLYQQLLHS
]	}		1		SLGAPGESTLLVRTSKLLVGLGLQLLVWLLL QTRSLLALQLHLTSSAPLLAAPTAVCSCSRCS
		1	1			APRSRCVARPAARTGLPTPAPASSPAPAASPA
İ				1		PAASPAPAESTA\PQPLILLPKP/PPAPGAPPPRP
		1	1			GAPPPRPAASPSPAASPAPPAASPVLTASPPLP
	Ì		1			AASPSPAASPAPPAASPVLTASPPLPAASPSPA
1	ł	ł	ł	l		ASPAPPAASPVLTASPPLPAASPALAASPVHT
	l	l		ĺ		ASPPVHVASPPVHTASPPVHVASPPVHTASPP
		ŀ	1	1		VHVASPPVHTASPHVHVASPPVHVASPPVHV
İ	1	ļ	1	ł		ASPPVHTASPPVHVASPPVHTASPHVHVASPP
		l				VHTASPPVHVASPPVHVASPPVHVAYPPVHV
						ASPPVHVASPPVHVASPPVSCSGDSTSDCFPP
424	1504	 	-			QPGAVFPHSLAPSLGGWSHLVAALP
434	1784	Α	3516	142	590	GGVNRPRSETEQVKTPVLISSWDYRHPPPRPA
		ļ				SFFVFLV*TGF\TALARMVLISWPCDLPTSASQ
		Ì				SAGITGVRHHA\RLLYFEQESHSVTQAGW\VQ
						WHNLGSLOPLSLEDRLSPGVLGCSALCRSGV
435	1785	A	3529	1 -	3161	RTKFGINMVTSRERGTTRLPKEG
	1,00	``	3327	•	5101	MSLVRAALEALDELDLFGVKGGPQSVIHVLA DEVQHCQSILNSLLPRASTSKEVDASLLSVVS
		•]	. !		FPAFAVEDSQLVELTKQEIITKLQGRYGCCRF
						LRDGYKTPKEDPNRLYY/ENPAELKLFENIEC
			1 1	,		EWPLFWTYFILDGVFSGNAEQVQEYKEALEA
		}	1 1			VLIKGKNGVPLLPELYSVPPDRVDEEYQNPHT
			1			VDRVPMGKLPHMWGQSLYILGSLMAEGFLA
			1			PGEIDPLNRRFSTVPKPDVVVQVYPSLPHGCS
			ł l	ł	1	SKSPSHQCTIISIRTTRKITAPVSILAETEEIKTIL
1						KDKGIYVETIAEVYPIRVQPARILSHIYSSLEIF
			}	1		LPFLNSVSGCNNRMKLSGRPYRHMGVLGTSK
					İ	LYDIRKTIFTFTPQFIDQQQFYLALDNKMIVE
}			!!	})	MLRTDLSYLCSRWRMTGQPTITFPISHSMLDE
			1 1	1		DGTSLNSSILAALRKMQDGYFGGARVQTGKL
,			1 [SEFLTTSCCTHLSFMDPGPEGKLYSEDYDDN
l]		YDYLESGNWMNDYDSTSHARCGDEVARYL DHLLAHTAPHPKLAPTSQKGGLDRFQAAVQT
				l	1	TCDLMSLVTKAKELHVQNVHMYLPTKLFQA
- [- [l	SRPSFNLLDSPHPRQENQVPSVRVEIHLPRDQ
						SGEVDFKALVLQLKETSSLQEQADILYMLYT
				ļ	Í	MKGPDWNTELYNERSATVRELLTELYGKVG
1					1	EIRHWGLIRYISGILRKKVEALDEACTDLLSH
ł	ł			l		QKHLTVGLPPEPREKTISAPLPYEALTOLIDEA
	-			ł	İ	SEGDMSISILTQEIMVYLAMYMRTQPGLFAE
	1			ľ	į	MFRLRIGLIIQVMATELAHSLRCSAEEATEGL
1	1	. 1			ļ	MNLSPSAMKNLLHHILSGKEFGVERSVRPTD
İ	İ	, i				SNVSPAISIHEIGAVGATKTERTGIMQLKSEIK
ļ						QSPGTSMTPSSGSFPSAYDQQSSKDSRQGQW
1					ŀ	QRRRRLDGALNRVPVGFYQKVWKVLQKCH
1	ŀ	٠	J			GLSVEGFVLPSSTTREMTPGEIKFSVHVESVL
l					ſ	NRVPQPEYRQLLVEAIL\VLTMLADIENHSIGS
				1		IIAVEKIVHIANDLFLQEQKTLGADDTMLAKD
				I		PASGICTLLYDSAPSGRFGTMTYLSKAAATY
436						VQEFLPHSICAMQ
	1786	A 1	3546	73	303	CDAL TWELL BUILT A BUILDING DODGE TO THE
- 1	1786	A	3546	73	393	CP*LTWELLEVKKAEVLQDSLDGRYSTPSSCL EQPDSCRPYGRSFYALEEKHVIFSLDVGETDN

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A:-Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible mucleotide deletion, \=possible nucleotide insertion KGKGKTIRGI*TFKGRKGGTYQREHDANPLA
437	1787	A	3554	5157	2939	PXSARSCWMRKG AVRAEPGLEELSSGLRAHSPSATTVCEPEAOG
				•		SASGCRYAAHPHWGLGGAAAAGGSWEPQPP RPVCEPAGRGKPHPPAAPRSPLLPGSRRRPHA AQPGARARTSPPPASARNMAARPAATLAWSL LLLSSALLREGCRARFVAERDSEDDGEEPVVF PESPLQSPTVLVAVLARNAAHTLPHFLGCLER LDYPKSRMAIWAATDHNVDNITEIFREWLK NVQRLYHYVEWRPMDEPESYPDEIGPKHWP TSRFAHVMKLRQAALRTAREKWSDYILFIDV DNFLTNPQTLNLLIAENKTIVAPMLESRGLYS NFWCGITPKGFYKRTPDY\VQIREWKRTGCFP VPMVHSTFLIDLRKEASDKLTFYPPHQDYTW TFDDIIVFAFSSRQAGIQMYLCNREHYGYLPIP LKPHQTLQEDIENLIHVQIEAMIDRPPMEPSQ YVSVVPKYPDKMGFDEIFMINLKRRKGQGGD RWLRTLYEQEIEVKIVEAVDGKALNTSQLKA LNIEMLPGYRDPYSSRPLTRGEIGCFLSHYSV WKEVIDRELEKTLVIEDDVRFEHQFKKKLMK LMDNIDQAQLDWELIYIGRKRMQVKEPEKA VPNVANLVEADYSYWTLGYVISLEGAQKLV GANPFGKMLPVDEFLPVMYNKHPVAEYKEY
						YESRDLKAFSAEPLLIYPTHYTGQPGYLSDTE TSTIWDNETVATDWDRTHAWKSRKQSRIYSN
438	1788	A	3563	130	527	AKNTEALPPPTSLDTVPSRDEL IFFNSSSLFCRVFCLFLRWSFTLVAQARVQ*C NLSSLQPLPPGFK*FSCLSPPRS*DYRRPPPRPA NFLYF**RQGFTVLGQAGLELLT/S/GDPPTSA SQSAGITGVSHRAWPVHAISTHISLVKTRPSLT TLG
439	1789	A	3565	446	1834	LLQPAMRKSPGLSDCLWAWILLSTLTGRSY GQPSLQDELKDNTTVFTRILDRLLDGYDNRL RPGLGERVTEVKTDIFVTSFGPVSDHDMEYTI DVFFRQSWKDERLKFKGPMTVLRLNNLMAS KIWTPDTFFHNGKKSVAHNMTMPNKLLRITE DGTLLYTMRLTVR\AECPMAFGRDFPM\D\AH ACPLKFGSYAYTRAEVVYEWTREPARSVVV AEDGSRLNQYDLLGQTVDSGIVQSSTGEYVV MTTHFHLKRKIGYFVIQTYLPCIMTVILSQVSF WLNRESVPARTVFGVTTVLTMTTLSISARNSL PKVAYATAMDWFIAVCYAFVFSALIEFATVN YFTKRGYAWDGKSVVPEKPKKVKDPLIKKN NTYAPTATSYTPNLARGDPGLATIAKSATIEP KEVKPETKPPEPKKTFNSVSKIDRLSRIAFPLL
440						FGIFNLVYWATYLNREPQLKAPTPHQ
440	1790	A	3568	1	350	STSSCFPAAAAAIMREIVHLQAGQCGNQIGAK FWEVISDEHGIDPTGTYHGDSDLQLERINVYY NEATGEAPVPSPTALRGPRGPCLG*RPPVPAG GKYVPRAVLVDMEPGTMDSV
441	1791	A	3569	2	1751	FVAVAGAVSGEPLVHWCTQQLRKTFGLDVS EEIIQYVLSIESAEEIREYVTDLLQGNEGKKGQ FIEELITKWQKNDQELISDPLQQCFKKDEILDG QKSGDHLKRGRKKGRNRQEVPAFTEPDTTAE VKTPFDLAKAQENSNSVKKKTKFVNLYTREG QDRLAVLLPGRHPCDCLGQKHKLINNCLICG RIVCEQEGSGPCLFCGTLVCTHEEQDILRGDS NKSQKLLKKLMSGVENSGKVDISTKDLLPH QELRIKSGLEKAIKHKDKLLEFDRTSIRRTQVI

NO: of nucle peptide cotide cotide cotide cotide sequence	SEQ ID	SEO ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
nucleotide seq- unice un							
Sequence 90,995 corresponding class tamino acid residue of peptide residue of peptide sequence s			1.00				
sequence uni		1 * *					
Part Part	1						
amino acid residue of sequence peptide sequence							
residue of peptide sequence Poptide Popti		1					
Peptide							
		ł	Ť			30420,00	
DDESDYFASDSNOWLSKLERETICKREEELK ELRHASRISKKYTIDFAGRKILEESKLAEYH SRLDETIOAJANGTI.NOPLIKLDESSEPELGYY SRLDETIOAJANGTI.NOPLIKLDESSEPELGYY VNPWMYOSPPQWVDHTGAASQKKAFRSGG GLEFNSFOHQLRIQDQEFQEGDGDGGCLSYH QPWASILLVRGIRAVEGRSWYTPHRGRLWIAA TAKKPSPQEVSELQATYRELIRKGWCLSYH QPWASILLVRGIRAVEGRSWYTPHRGRLWIAA TAKKPSPQEVSELQATYRELIRKGWCLSYH QGAKKCLMKQNKAV 442 1792 A 3576 1 2019 MPRSHTIGERLGEKEGSQCAENFSPNI.SVTK KTAGVKPYECTICGRAFMRI.SSLTRHMRSHT ARANERYYKCKEGGRAFSISQIS.KWIERSH TGEKPYKCKQCGKI.FIYHQPFQRHERTHIGER PYECK,GCGKA.FICSSIR.WIERSHTGERYPEC KQCGKA.FICSSSIR.WIERHTGERYPECK GKAFISTSVLTHMITHORPYYKCKEGGKA FIFPSELRVHERHTGERYPECK GKAFISTSVLTHMITHORPYYKCKEGGKAFRNITHIGERY HVOXITIGERYYKCKEGGKAFRSTYPHHERT HTGGEPYECKCGKTFNYFLDLKHIKRNHTG ERYYECKCECKTFSTSTYPHHERT HTGGEPYECKCGKTFNYFLDLKHIKRNHTG ERYPECKCECKTFSTSTYPHHERT HTGGEPYECKCGKTFNYFLDLKHIKRNHTG ERYPECKCECKTFSTSTYPHHERT HTGGEPYECKCGKAFRSTSYTRHERT HTGGEPYECKCGKTFNYFLDLKHIKRNHTG ERYPECKCECKTFSTSTYPHHERT HTGGEPYECKCGKAFSSTSTYRHERT HTGGEPYECKCGKAFSSTSTYRHRTGGRYPECKEGGK AFITYTSFQGHMRMITGERYYKCKECGKAFS LISSFRRATIRIINFYRKPLECYQCGCAFSVSTS LKSPRRATIRIINFYRKPLECYQCGCAFSVSTS LKSPRRATIRIINFYRKPLECYQCGCAFSVSTS LKSPRRATIRIINFYRLECYQCGAFSVSTS LKSPRRATIRIINFYRLECYQCGAFSVSTS LKSPRRATIRIINFYRLECYQCGAFSVSTS LKSPRRATIRIINFYRLECYQCGAFSVSTS LKSPRRATIRIINFYRLECYQCGAFSVSTS LKSPRRATIRIINFYRLECYQCGAFSVSTS LKSPRRATIRIINFYRLECYQCGAFSVSTS LKSPRRATIRIINFYRLECYQCGAFSVSTS LKSPRRATIRIINFYRLECYQCGAFSVSTS LKSPRRATIRIINFYRLECYQCGAFSVSTS LKSPRRATIRIINFYRLECYQCGAFSVSTS LKSPRRATIRIINFYRLECYQCGAFSVSTS LKSPRRATIRIINFYRLECYQCGAFSVSTS LKSPRRATIRIINFYRLATICGAFSVTY VAVVLYPH-STLFACKLGWCKAALASVV-IVQ DSGRGRWLTPVIPALWEAKAGGSRGQEKTIL ANTVEPH-JY DSGRGRWLTPVIPALWEAKAGGSRGGEKTIL ANTVEPH-JY DSGRGRWLTPVIPALWEAKAGGSRGGEKTIL ANTVEPH-JY HTMYSVLAGGSFRATISTCHAWAGGSFRATISTCHAWAGG FURDINGDALLDSGCCC WERRIGAFGGER AFTEGERERAYSTCHAWAGGSFRATISTCHAWAGG FURDINGDALLDSGCCC WERRIGAFGGER AFTEGERERAYSTCHAWAGGFRATISTCHAWAGG KKMMINEGELTARGARPSW KAAK TGGEREGAFASSFLOGALSWFRERPA AISIENSIDDTSNFFEOTWEHEREPA AISIENSIDDTSNFFEOTWEHERPA AISIENSIDDTSNFFEOTWEHERPA AISIENSID	ĺ	1				ł	
ELRHASRISKKVTIDFAGRKILEENSLAEVIH SRLDETIQIAJANGTLAOPLTKI DRSSEPEIGVI VNPNMYGSPQWVDHTGAASQKKARSSSEP GLEFNSFOHQLARGDQETQEGDEGGEGGGGCCLSVH QPWASILLYRGIKRVEGRSWYTPHRGRL WYAA TAKKPSPQEVSELQATYRLIRGKOWCLSVH PSGCLLGCVDLIDCLSOKORKDQPDISQESDE PFVFICKNPQEMVYEPIKGDPKIWKLDSKIH QGAKKGLMKQNKAV QGAKKGLMKQNKAV QGAKKGLMKQNKAV ARABERYKCKECGGRAFISSQLSKHERSH ARABERYKCKECGGRAFISSQLSKHERSH TGEKPYCKCQGGKAFWARSSLTRHMESHT ARABERYKCKECGGRAFISSQLSKHERSH TGEKPYCKCKGGKIFSPHALSTTK KTAGVKPYECTICGEAFMARSSLTRHMESHT ARABERYKCKECGGKAFSISQLSKHERSH TGEKPYCKCKGGKIFSPHACKECGG GKAFISTISVLTHMITHNORPYKCKECGGA FIPSFLEVHERHITGERPYYCKCEGGKAFRW TSIQHERIHTGEKPYKCKECGKAFRW TSIQHERIHTGEKPYKCKECGKAFRW TSIQHERIHTGEKPYKCKECGKAFRW TSIQHERIHTGEKPYKCKECGKAFRW TSIQHERIHTGEKPYKCKECGGKAFRW TSIQHERIHTGEKPYKCKECGKAFRW TSIQHERIHTGEKPYKCKECGKAFRW TSIQHERIHTGEKPYKCKECGKAFRW TSIQHERIHTGEKPYKCKECGKAFRW TSIQHERIHTGEKPYKCKECGKAFRW TSIQHERIHTGEKPYKCKECGKAFRW TSIQHERIHTGEKPYKCKECGKAFW TSIQHAMA TSIGHTATA TSIQHERIHTGEKPYKCKECGKAFW TSIQHAMATATA TSIQHAMATATA TSIG		 	 	 	sequence		
SRLDETIQALANGTLNOPLTKLDRSSEEPLGAV VNPNMYGSPPQWVDTHGAASQKARAPRSGP GLEFNSCHOLRICDOGEPCGGPGGGWCLSVM OPWASLLVRGIKBVEGGSWYLTPRIGHERI WTAA TAKKPSPQEVSEL QATVELLRGKDVEFPNDY PSGCLLGCVDLIDCLSQKOPKLDGFDISQSSD PFVFICKNPQEMVVKFPIKGNPKIWKLDSKIH QGAKKGLMKQNRAV 442 1792 A 3576 1 2019 MPRSHTIGERLEGKEGSQCAENFSPNLSVTK KTAGVKPYECTICGKAFMRLSSLTRHMRSHT ARANERYYKCKECGKAFSISQLSKUERSH TGEKPYKCKQCGKTFIYHQPFQRHERTHIGER PYECKQCGKALSCSSLSWIPERHTIGERYPYCK KQGKAFSCSSIRVHERITIGERYPYCKGCGKAFSISQLSKUERSH TGEKPYKCKQCGKAFSISSIRVHERITIGERYPYCK KQGKAFSCSSIRVHERITIGERYPYCKGCGKAFSFORT PYECKQCKALSCSSLSWIPERHTIGERYPYCKGCGKAFSFORT FFSFLRVHERITIGERPYCKCCGCKAFRNY TSIQHERHTIGEKPYKCKECGKAFPAFRAY HVRYHTGERPYKCKECGKAFSPNSTYRHIEBT HTGGEFYPCKKCKFTNYPDLIKHKRNHTO EKPYECKECAKTFISLENFRRHMITHTGBCPYECKC CGKAFSSSYIRHRRTHTGEKPYECKCGC AFIFTSFLRVHERTHTGEKPYECKCGC AFIFTSFLRVHERTHTGEKPYECKCGC AFIFTSFLRVHERTHTGEKPYECKCGC AFIFTSFLRVHERTHTGEKPYECKCGC CGKAFSSSYIRHRRTHTGGKPYECKCGC AFIFTSFLRVHERTHTGGKPYECKCGC CGKAFSSSYIRHRRTHTGGKPYECKCGC AFIFTSFLRVHERTHTGGKPYECKCGCAFASYAT ALSEMSTONTSFLRVHERTHTGGGKPYECKCGCAFASYAT ALSEMSTONTSFLRVHERTHTGGGKPYECKCGCAFASYAT ALSEMSTONTSFLRVHERTHTGGKPYECKCGAFASYAT ALSEMSTONTSFLRVHERTHTGGKPYECKCGAFASYAT ALSEMSTONTSFLRVHERTHTATHTGGKPYECKCGCAFASYAT ALSEMSTONTSFLRVHERTHTATHTGGKPYECCC A		į					
WNPMMYGSPQWVDHTGAASQKKAPRSSGG GLEFNSPGOLGLOGDETQEGDGGGGCLSVH QPWASILLYRGIKRVEGRSWYTPHRGRL WMA TAKKPSPQEVSELQATYRLIRGKOVEPNDY PSGCLLGCVDLIDCLSQKQFKEQFPDISQESDE PFVFICKNYGEWYKFPKGWFKIWKLDSKH PSGCLLGCVDLIDCLSQKQFKEQFPDISQESDE PFVFICKNYGEWYKFPKGWFKIWKLDSKH ARABEKPYKCKEGGAFSISQLSKYHERSH TGKRYYPKCKCGGKAFKRISSLTRHMRSHT ARABEKPYKCKEGGAFSISQLSKYHERSH TGKRYYKCKEGGKAFKRISSLTRHMRSHT ARABEKPYKCKEGGAFSISQLSKYHERSH TGKRYYKCKCGGAFSISQLSKYHERSH TGKRYYKCKCGGAFSISQLSKYHERSH TGKRYYKCKCGGAFSISQLSKYHERSH TGKRYYKCKCGGAFSISQLSKYHERSH TGKRYYKCKCGGAFSISQLSKYHERSH TGKRYYKCKCGGAFSISQLSKYHERSH TGKRYYKCKCGGAFSISQLSKYHERSH TGKRYYKCKCGGAFSISQLSKYHERSH TGKRYYKCKCGGAFSISTSTHHERSH TGKRYYKCKCGGAFSISTSTHHERSH TGKRYYKCKCGGAFSISTSTHHERSH TGKRYYKCKCGGAFSISTSTHHERSH TGKRYYKCKCGGAFSISTSTHHERSH TGKRYYKCKCGGAFSISTSTHHERSH TGKRYYKCKCGGAFSISTSTHHERSH TGKRYYKCKCGGAFSISTSTHHERSH TGKRYYKCKCGGAFSISTSTHHERSH TGKRYYKCKCGCGAFSISTSTHHERSH TGKRYYKCKCGCGAFSISTSTHHERSH TGKRYYKCKCGCGAFSISTSTHHERSH TGKRYYKCKCGCGAFSISTSTHHERSH TGKRYYKCKCGCGAFSISTSTHHERSH TGKRYYKCKCGCGAFSISTSTHHERSH TKKRYYKCGCCGAFSISTSTHHERSH TKKRYYKCGCCGAFSISTSTHHERSH TKKRYYKCGCCGAFSISTSTHHERSH TKKRYYKCGCCGAFSISTSTHHERSH TKKRYYKCGCCGAFSISTSTHHERSH TKKRYYKCGCCGAFSISTSTHHERSH TKKRYYKCGCCGAFSISTSTHHERSH TKKRYYKCGCCGAFSISTSTHHERSH TKKRYYKCGCCGAFSISTSTHHERSH TKKRYYKCGCCGAFSISTHHERSH TKKRYYKCGCCGAFSISTHHERSH TKKRYYKCGCCGAFSISTHHINTSH TKKRYYKCGCCGAFSISTHHINTSH TKKRYYKCGCCGAFSISTHHINTSH TKKRYYKCGCCGAFSISTHHINTSH TKKRYYKCGCCGAFSISTHHINTSH TKKRYYKCGCCGAFSISTH TKKRYYKCGCCGAFSISTH TKKRYYKCGCCGAFSISTH TKKRYYKCGCCGAFSISTH TKKRYYKCGCCCGAFSISTH TKKRYYKCGCCCGAFSISTH TKKRYYKCGCCCGAFSISTH TKKRYYKCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		[l	CDI DETIONANCEI NODI TVI DECERDI CIE
GLEFNSPOHOLRIQDQEFQEGFDGGWCLSVM OPWASILVRIGIKVEGGRSWYTPHRIGILWIAA TAKKPSPQEVSELQATYRLLRGKDVEPPNDY PSGCLIGCVDLDGLSQKQPKDGPFDISQESD PFVFICKNPQEMVYKFPIKGNPKINKLDSKIH QGAKKGLMKQNKAV 442 1792 A 3576 1 2019 MPRSHTGERLEGKEGSQCAENFSPNLSVTK KTAGVKPYECTICGKAFMRLSSLTRHMRSHT AIRANEKPYKCKEGGAFSISQIS,KWIERSH TGEKPYKCKQGGKIFYHQPFQRHBRTHIGEK PYECKQGGKALSCSSIR.WIERSHTGEKPYECK GGAKAFSCSSIR.WIERHTGEKPYECKQGKAFS TSIQHERIHTGEKPYKCKECGKAFS TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYECKECGK AFFYPTSPQGHMRMHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYECKECGK AFFYPTSPQGHMRMHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYECKECGK AFFYPTSPQGHMRMHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYECKECGK AFFYPTSPQGHMRMHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYECKECGK AFFYPTSPQGHMRMHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYECKECGK AFFYPTSPQGHMRMHTGEKPYKCKECGKAFW THAWAY TSIQHERIHTGEKPYECKECGK AFFYPTSPQGHMRMHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYECKECGK AFFYPTSPQGFMRMHTGEKPYKCKECGKAFSW TSIQHERIHTGEKPYECKECGKAFSW TSIQHERIHTGEKPYECKECGKAFSW THEKTTER THOTE THINDH TO THE TOT TO THE TOT TO THE TOT T		1					WIDNIA CORDONALDUTE A CORRESECTE OF
OPWASILVRGIKRVEGRSWYTPHRGRILWIAN TAKKPSPQEVEGOPTPOLOGYPHONY FSGCLLGCVDLIDCLSQKOFKEQPPOISQESDE PFVFICKNPQEMVVKFPIKOPKIWKLDKIM GGAKKGLMKQNKAV OGAKKGLMKQNKAV OGAKKGLMKQNKAV MPRSHTIGERLGEGREGSQCAENFSPNLSVTK KTAGVKFYECTICKAFMRLSSLTRHMRSHT AIRANEKPYKCKEGGRAFSLSQLSKUERRHTGER FOR STANKEN TORKYYKCKEGGRAFSLSQLSKUERRHTGER FYECKQCGKALSCSSLRVHERHTGER FYECKQCGKALSCSSLRVHERHTGER FYECKGCGKALSCSSLRVHERHTGER FYECKGCGKALSCSSLRVHERHTGER FYECKGCGKALSCSSLRVHERHTGER FYECKGCGKALSCSSLRVHERHTGER FYECKGCGKALSCSSLRVHERHTGER FYECKGCGKALSCSSLRVHERHTGER FYECKGCGKALSCSSLRVHERHTGER FYECKGCGKATSRLST FYECKGCGKALST FYECKGCGKALST FYECKGCGKALST FYECKGCGKATSRLST FYECKGCGCT FYECKGCGCKATSRLST FYECKGCGCT FYECKGCGCT FYECKGCGCT FYECKGCGCT FYECKGCGCT FYECKGCGCT FYECKGCGCT FYECKGCGCT FYECKGCGCT FYECKGCGCT FYECKGCGCT FYECKGCGCT FYECKGCGT FYECKGCGT FY						j	
TAKKPSPGEVSELQATYRILRGKDVEPPNOF PSGCLIGCVDLIDCLSQKOFKLOPPDISQESDS PFVFICKNPQEMVVKPJEKGNPKIWKLDSKIH QCAKKGIMKQNKAV 442 1792 A 3576 1 2019 MPRSHTGERLEGKEGSQCAENFSPNLSVTK KTAGVKPYECTICGKAFMELSLIRKHMENTT AIRANEKPYKCKEGGRAFISJGSKJERSH TGEKPYKCKQCGKATISHOPPNETRTHAGE PYECKQCGKALSCSSIRVHERHITGEKPYACKUG GKARISTISLKHMENSTT TGEKPYKCKQCGKTPIYHQPPORHERTHAGE PYECKQCGKAISCSSIRVHERHITGEKPYACKUG GKARISTISLYHMINHORDPYKCKEGCKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYKCKQCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYFRHERT HTGEKPYECKGCGKAFRSYTR LKKPMRNAQSDRKIJVKCEK-EKVPNSNFR LKKPMRNAQSDRKIJVKCEK-EKVPNSNFR LKKPMRNAQSDRKIJVKCEK-EKVPNSNFR LKKPMRNAQSDRKIJVKCEK-EKVPNSNFR LKKPMRNAQSDRKIJVKCEK-EKVPNSNFR QSCENSHREKSCCCK-YNKROKAFSVST LKKPMRNAQSDRKIJVKCEK-EKVPNSNFR LKKPMRNAQSDRKIJVKCEK-EKVPNSNFR QSCENSHREKSCCCK-YNKROKAFSVST LKKPMRNAQSDRKIJVKCEK-EKVPNSNFR LKKPMRNAQSDRKIJVKCEK-EKVPNSNFR LKKPMRNAQSDRKIJVKCEK-EKVPNSNFR LKKPMRNAQSDRKIJVKCEK-EKVPNSNFR LKKPMRNAQSDRKIJVKCEK-EKVPNSNFR LKKPMRNAQSDRKIJVKCEK-EKVPNSNFR LKKPMRNAQSDRKIJVKCEK-EKVPNSNFR LKKPMRNAQAFSPTOSMSNHTKERYT MTKYTLENFSNLIAQHEREBMRQKKILEKV MEEGGLKDEKRLRKV MEEGGLKDEKRLRKV MEEGGLKDEKRLRKSAHARKETEFLIKT RLCLEDFISLKVIGKGAFGEVRIJVALDSHQL GFHRDIKPNNIJKRFDFTQMNISKRK MEEGGLKDEKRKGKLEKV MEEGGLKDEKRKGKLEKV MEEGGLKDEKRLRKV MEEGGLKDEKSKGRUTSLURFCCE WERRIGAFOVERSTORT KKVMMVERTLIPPSVIJKHEGLILRFCCE WERRIGAFOVERSTORT KKVMMVERTLIPPSVIJKHEGLILRFCCE WERRIGAFOVERSTORT KKVMMVERTLIPPSVIJKHEGLILRFCCE WERRIGAFOVERSTORT KKVMMVERTLIPPSVIJKHEGLILRFCCE WERRIGAFOVERSTORT KKVMMVERTLIPPSVIJKHEGLILRFCCE WERRIGAFOVERSTORT KKVMMVERTLIPPSVIJKHEGLILRFCCE WERRIGAFOVERSTORT KKVMMVERTLIPPSVIJKHEGLILRFCCE WERRIGAFOVERSTORT KKMMTAGAFOVERSTORT KKVMMVERTLIPPSVIJKHEGLILRFCCE WERRIGAFOVERSTO		l]	,		ł	
PSGCLLGCVDLIDCLSQKQFKLQCPPDISQESDS PFVFICKNPQEMVVKPHKGPKINPKINVKLDSKIH QGAKKGLMKQNKAV		1	l				TAVVECENCE CATVELLE CVENTER TO
1792 A 3576 1 2019 MPRSHTGERLCEGKEGSOCAENFSENLSYL A 3576 1 2019 MPRSHTGERLCEGKEGSOCAENFSENLSYL KTAGVKPYECTICKAFMRLSSLTREMRSHT AIRAVEKPYKCKECGRAFSSQLSKVIERSH TGEKPYKCKQCGKLSTYLOPHOPORHERTHERSHT AIRAVEKPYKCKECGRAFSSQLSKVIERSH TGEKPYKCKQCGKTTYHOPPORHERTHERSHT AIRAVEKPYKCKECGKAFSSSSSLRVHERHTGEKPYECKCGKAFSTSSSSSRVIERHTGEKPYECKECGKAFISTTSVLTHMITHNORPYKCKECGKAFISTSTYLTHMITHNORPYKCKECGKAFISTSTYLTHMITHNORPYKCKECGKAFISTSTYLHIERTHTGEKPYECKCGKAFSTSTSPHIERTHTGEKPYECKCGKAFSTSSTFHIERTHTGEKPYECKCGCKAFSTSTSTHIERTHTGEKPYECKCGCCCGKAFSCSSSTRWHERHTHTGEKPYECKCGCCCGKAFSCSSSTRWHERHTHTGEKPYECKCGCCCCGKAFSCSSSTRWHERHTHTGEKPYECKCGCCCCGKAFSCSSSTRWHERHTHTGEKPYECKCGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC							PSGCLI COMPLINGLEON OF TRANSPORTED TO THE PROPERTY OF THE PROP
QGAKKGLMKONTAV			ŀ				PENELCY IDOE AND CENTRAL POPULATION OF THE PROPERTY POPULATION OF THE PROPE
442 1792 A 3576 1 2019 MPRSHTGERLCEGKEGSQCAENFSPISJYTK KTAGVKPYECTICGKAPMRLSSLTRHMRSHT AIRANEKPYKCKECGAFSLSQILSKNIERSHT GEKPYKCKQCGKTFIYHQPFQRIEBERHIGEK PYECKQCGKALSCSSSLRVHERIHTGEKPYACKEC GGAFISJTSULTHMITHSOPPKYKEKCGKAFWS TSIQIHERIHTGEKPYKCKCGGKAFRWS TSIQIHERIHTGEKPYKCKCCGKAFRWS TSIQIHERIHTGEKPYKCKCCGKAFRWS TSIQIHERIHTGEKPYKCKCCGKAFRWS TSIQIHERIHTGEKPYKCKCCGKAFRWS TSIQIHERIHTGEKPYKCKCCGKAFRWS TSIQIHERIHTGEKPYKCKCCGKAFRWS TSIQIHERIHTGEKPYKCKCCGKAFRWS TSIQIHERIHTGEKPYKCKCCGKAFRWS TSIQIHERIHTGEKPYKCKCCGKAFRWS TSIQIHERIHTGEKPYKCKCCGKAFRWSTSIK HYPTSFQGHMRMHTGEKPYKCKECGKAFS LHSSFRRHTTRINNYRKPLEC*QICCKAFVSTS LKKFMRNAQSPKLKYKCKECGKAFS LHSSFRRHTTRINNYRKPLEC*QICCKAFVSTS LKKFMRNAQSPKLKYCKECKAFVSTS LKKFMRNAQSPKLKYCKECKAFVSTS LKKFMRNAQSPKLKYCKECKAFVSTS LKKFMRNAQSPKLKYCKECKAFVSTS VAVVTPCSTLEKCLWCWCKRAALSVV/IVQ DSGGRW-LTPIPPALWEAKAGGSGGQEIKTL ANTVKPHLY 443 1793 A 3578 287 114 DFYERKFEQFEGHKQIVNKWRDLLCSWKK LSIEKKSVLQNNL*PSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFAAMAMTGSTFCSSMSNHTKERVT MIKVTLENFYSNLAQHEERGKKLEKV MEEGLADEEKLVGRAAFECVELVGKKLEKV MEEGLADEEKLVGRAAFECVELVGKKLEKV MEEGLADEEKLVGRAAFECVELVGKKLEKV MEEGLADEEKLVGRAAFECVELVGKKLEKV MEEGLADEEKLVGRAAFECVELVGKKLEKV MEEGLADEEKLVGRAAFECVELVGKKLIGH VYAMKLIKAADMEKECVGHIRAERDILVFA DSLWVVKMFYSFQDKLNLVLIMEFLPGDM MTLLMKKDTLTEETCFYTVALDSHIQL GFHRDIKPDNLLLDSKFKALDLLRFCCC WEHRIGAPGVEEIKSNSFFEGVDWEHIRRPA AISEIKSIDDTSFPEFESDTVALDSHIGH GFHRDIKPDNLLLDSKKAKLLIHFCCC WEHRIGAPGVEEIKSNSFFEGVDWEHIRRPA AISEIKSIDDTSFPEFESDTVALDSHIGL GFHRDIKPDNLLLDSKRAALDHERGC WEHRIGAPGVEEIKSNSFFEGVDWEHIRRPA AISEIKSIDDTSFPEFESDTVALDSHIGH CKVMMWKETLTFPEVPVISEKAADLLIHFCCC WEHRIGAPGVEEIKSNSFFEGVDWEHIRRPA AISEIKSIDDTSFPEFESDTAMABELLE TOYKNKDWYFNYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTGGEKRFAYSFLQQLRYVDEAHQYILEPD GGSRGKGEHPYVEQEIKFFAKVVLLDQYFK NHRIYLFISAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGRDATSIVNCLHLIQOT LDAATVMKTGLESVYKSALRAALGECLAAFAGA KTMENLKQGGFTHTRNQPKGVTQINYTTVA LLPMLSSLFEHIGOHQFGEDLLIEDVGVSCYRI LTSLYVALGTSSAALGECLAAFAGA	i	ļ	1				
TAGOVEPYECTICGKAPMRLSSLTRHMRSHT	442	1700	<u> </u>	3576		2010	
AIRALEKPYKCKECGRASLSQULSKIERSH TGEKPYKCKQCGKTIFIYHOPPGRERTHIGER PYECKQCGKALSCSSLRVHERIHTGEKPYRCKECGK KQCGKATSCSSSIKVHERIHTGEKPYRCKECGK GKAFISITSVLTHMITHINGPKYPKCKECGKAFRW STSIGHERIHTGEKPYKCKECGKAFRW STSIGHERIHTGEKPYKCKECGKAFRW STSIGHERIHTGEKPYKCKECGKAFRW STSIGHERIHTGEKPYKCKECGKAFRW STSIGHERIHTGEKPYKCKECGKAFRW STSIGHERIHTGEKPYKCKECGKAFRW HNRVHTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT HTGEKPYECKECGKAFSRISYFRHIERT RIGKAFSRIGHTSTRICKTERT RIGKAFSRIGHTSTRICKTERT RIGKAFSRIGHTSTRICKTERT RIGKSCQCK*YFRKDTR*FMYSQV HPHNIVSNSHOPYRICGSBFRILYTN*INSINSPIT ANTVKPHLY ANTVKPHLY ANTVKPHLY DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK LSIKKSVLQNNL*PSAASMRQKLEKV MEEGGLKDEEKRLRRSAHARKETEFRLKTT RIGKEPSESLKVIGRAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKINIVLIIMEPLPGGDM MTLLMKKDTLTEETCPYTALADISHQL GFHRDIKPDNLLIDSKGHVKLSDFGLTGIK KAHTEFYRNINHSLPSDFTPGMINSKRKAE TWKRNRRQLAFSTVGTPDVIAPEVFMQTGYN KLCDWSLGVIMYEMLIGYPFFCATSHNEY KKVMWKETLTFPPEVPISKAKADLIRFCCE WEHRIGAPGVEEKSNSFFEGVDWEHIRRPA AISIEKSIDDTSNFDEFPESDILRFTVATSNHIP TDYYKNKDWFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 I 6169 RTRGIEKRFAYSFLQQLRYVDEAHQYILEFD GGSRGKGEHPYVEQIKIFAKVVLLIDGYTK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGONATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDMAAEDLE KTMENLKQGGFTHTRNQPKGVTQINNYTTVA LLPMLSSLFEHIGQHGFGEDLLIEDVYCSYRI LTSLYALGTSKSALIGECLAAFAGA LTPMLSKSLPCTSGRAFALDMAAEDLE KTMENLKQGGFTHTRNQPKGVTQINNYTTVA LLPMLSSLFEHIGQHGFGEDLLIEDVYCSYRI LTSLYALGTSKSALIGECLAAFAGA LTPMLSSLFEHIGOHGFGEDLLIEDVYCSYRI LTSLYALGTSKSALIGECLAAFAGA	442	1792] A	33/6	1	2019	MPRSH1GERLCEGREGSQCAENFSPNLSVTK
TGEKPYKCKQCGKTFINHQPFQBERTHIGER PYYECKQCGKALSCSSLRVHERIHTGEKPYEC KQCGKAFSCSSIRVHERIHTGEKPYACKEC GKAFISITISVLTHMITHNGDRPYKCKECGKAF FIPPSILRVHERIHTGEKPYKCKQCGKAFRWS TSIQHERIHTGEKPYKCKQCGKAFRWS TSIQHERIHTGEKPYKCKECGKAFRWS TSIQHERIHTGEKPYKCKECGKAFRWS TSIQHERIHTGEKPYKCKECGKAFRWS TSIQHERIHTGEKPYKCKECGKAFRWS TSIQHERIHTGEKPYKCKECGKAFRWS TSIQHERIHTGEKPYKCKECGKAFRWS CKDCGKVKIFPSALRTHERTHTIGEKPYECKECGK AFFORTSTOCHTSILENFRRHMITHTGDGPY KCRDCGKVFIFPSALRTHERTHTIGEKPYECKECGK AFFORTSTOCHTSILENFRRHMITHTGDGPY KCRDCGKVFIFPSALRTHERTHTIGEKPYECKECGK AFFORTSTOCHTSILENFRRHMITHTGDGPY KCRDCGKVFIFPSALRTHERTHTIGEKPYECKECGK AFFORTSTOCHTSILENFRRHMITHTGDGPY KCRDCGKVFIFPSALRTHERTHTIGEKPYECKECGKAFS LHSSFRWHTRIHNYEKPLEC*QICKAFSSVSTS LKKPMRNAQSDRKLYRCER*EVKPKCKCCKAFS LHSSFRWHTRIHNYEKPLEC*QICKAFSSVSTS LKKPMRNAQSDRKLYRCER*EVKPNSNCF QSCENSH*PREKSQQCK*YRKRDTR*PMYSQV PHNHVSVSNGPYRCGSPRLYNT*NISNRNL VAVVTP*SCTLFKCLWCKRAALSVV*NVQ DSGRGWLTPVIPALWEAKAGGSRGQEIKTIL ANTVKPHLY DSGRGWLTPVIPALWEAKAGGSRGQEIKTIL ANTVKPHLY 443 1793 A 3578 287 114 DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK LSIIKKSVLQNNL*PSSALRMFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAMAMTGSPTCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKLEKV MEEGGLDDEEKRLRSAHARKEFLRLKT RLGLEDFESIKVIGRGAFGEVRLVQKKDTGH VYAMKLRKADMLEKEQVGHRARRDILVGK MEKENGLAFSTVGTPDYIAPEVFMQTIGYN KLGDWWSKFYSQDKLNYLIMEPLPGDA MTLLMKKDTLTEETQFYIAETVLAIDSHQL GFHRDIKPDNLLLDSKGHVKLSDFICONMNSKRAE TVKRNIRQLAFSTVGTPDYIAPEVFMQTIGYN KLCDWWSLGVMYPSGCDLTDJFTQNMNSKRAE TVKRNIRQLAFSTVGTPDYIAPEVFMQTIGYN KKVMNWKETLTPPEVPISEKAKDLIRFCCE WEHRIGAFGVEEKSNSFFEGVDWHIRERPA AISIEKSIDDTSNDDEFPSDULKPTVATSNHTE TDYXNKDWFNYTYKREGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKFFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGBHFFYQEIKFFAKVYLDLDQYFK NHRLYFILSAASRFLCSGGHASNKEKEMYTSL FCKLGVLVRHIBSLGONDATSIVVACHLILGQT LDARTVMKTGLESVKSALRAFLDNAASDLE KTMENLKQGGFTHTRNQPKGVTQINYTTVA LLPMLSSLFEHIGGRQFGEDLLEDQVGSCYRI LTSLYALGTSKSIYVERGRSALGECLAAFAGA	1	1					
PYECKQCGKALSCSSSLRVHERIHTGEKPYACKIEC KQCGKAFSCSSSIRVHERIHTGEKPYACKIEC GKAFISTTSVLTHMITHNGDRPYKCKECGKAFRUS TSIQHERIHTGEKPYKCKQCGKAFRUS TSIQHERIHTGEKPYKCKQCGKAFRUS TSIQHERIHTGEKPYKCKECGKAFRUSYPRIHERT HTGEKPYECKECKTFNYPLDLKIHKRNHTG EKPYECKECAKTFISLENFRHIMTHTGDGPY KCRDCGKVFIFPSALRTHERTHTGEKPYECK CCKAFSCSSYRHKRTHTGEKPYECKC CCKAFSCSSYRHKRTHTGEKPYECKC CCKAFSCSSYRHKRTHTGEKPYECKECK AFVPTSFQGHMRMHTGEKPYKCKECGKAFSTS LKKPMRNAGSDRKLYKCEK*EKVPNSNRCF QSCENSH*REKSCQCK*YRKRDTR*PMYSQV PHNHYSVSNGPPR/CCSPHLYNT*NSINRNL VAAVTP*CSTLFKCLWCWCKRAALSVV*/IVQ DSGRGRWLTPVTALWEAKAGGSGEIKTIL ANTVKPHLY 443 1793 A 3578 287 114 DPYERKFEQFIEGHKQIVNKWRDLLCSWKRK LSIIKKSVLQNNL*PSAASMEFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFAAMAMTGSTPCSSMSNHTKERVT MIKVTLENFYSNLAQHEEREMRQKKLEKV MEEGGIKDEEKRLRRSAAMERGVLFF 445 1795 A 3584 1 6169 GFRDJKPONLLDSKGKLIKSV KAHRTEFYRNLNHSLPSDFTFQNMNSKRAE TWKRNRQLAFSTVGTPDYLAFEVRATSHHE TDYKNKDWYENTYTKREGLTARGAFSYM KAAK 445 1795 A 3584 1 6169 RTRGEEKRAYSFLQQLIRYYDEAHUTLBHQL LDARTVMKTGLSFGGALSHAFLER KAAK 445 1795 A 3584 1 6169 RTRGEEKRAYSFLQQLIRYYDEAHUTLBHQL LDARTVMKTGLSFGGALSHAFLER LDARTVMKTGLSFGGALSHKEKEWTYSL FCKLGVLVNRHSILFGNDATSIVACHLILGQT LDARTVMKTGLSFGAKSALRAFLDNAABDLE KTMENLKQQGFTHTRNQPKGVTQINYTTVA LLPMLSSLFEHIGQHQFGEDLLLEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA			١ ،				
KQCGKAFSCSSIRVHERTHTGEKPYACKEG GKAFISITSULTIMITHEORPYKCKQCGKAFRWS TSIQHERHTGEKPYKCKQCGKAFRWS TSIQHERHTGEKPYKCKEGKSSARPAFRY HVRVHTGERPYKCKEGKAFSRISYFRIHERT HTGEKPYECKKGKTFNYPLDLKHKRNHTG EKPYECKKGKTFNYPLDLKHKRNHTG EKPYECKKGKTFNYPLDLKHKRNHTG EKPYECKKGKTFNYPLDLKHKRNHTG EKPYECKKGKTFNYPLDLKHKRNHTG EKPYECKKGKTFNYPLDLKHKRNHTG EKPYECKKGKTFNYPLDLKHKRNHTG EKPYECKKGKAFSSISTRIHERTHTGEKPYECKCGKAFS AFIVPTSFQGHMRMHTGEKPYECKEGKAFS LHSSFRRHTRIHNYEKPLEC*QICGKAFSVSTS LKKPMRNAQSDRKLI YKE*EK*EXVPNIVQ DSGRGRWLTPVPALWEAKAGGSRQEIKTIL ANTVEPHLY 443 1793 A 3578 287 114 DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK LSIIKKSVLQNNL*PSAASMEFQKVFF 444 1794 A 3582 3335 1909 HLFFSLLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLAQHEEREMRQKKLEKV MEEGGKDEKRKRSAHARKETEFLRKAT RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKLRKADMLEKEQVCHRAARDLLVEA DSLWVVKMYSPQDKLNYLMEPLFQGDM MTLLMKKDTLTEETGPYIAETVLAIDSHQL GFHRDIKDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPFVRQTGYM KLCDWWSLGVIMYPSGUNGHUGYPFFCSETPQETY KKVMNWKETLTPPEVPISEKAKDLURFCCE WERRIGAPGVEEKSNSFFEGVDWEHIRERPA AISEEKSIDDTSNTDEFPESDILRPTVATSNHTE TDYKNKDWYFNYTYKRFEGLTARGAPSYM KAAK 445 1795 A 3584 1 6169 RTRGEKRFAYSFLQQLRYYDEAGHYLLEFD GGSRGKGEHFPYFQEIKFFAKVVLPLIDQYTF NRRYFISAASRFLLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVVACHLILGOT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGGFTHRNQPKGVTQINYTTVA LLPMLSSLFEHIGQHQFGEDLLLEDVQVSCYRI LTSLYALGTSKSIYVERGRSALGFECLAAFAGA		1					
GKAFISTTSVLTHMITHNGDRPYKCKECGKAFRWS TISIGHERHITGEKPYKCKCGCKAFRWS TISIGHERHITGEKPYKCKECGKSFSARPAFRV HVRVHTGEKPYKCKECGKAFSRISYFRIHERT HTGEKPYECKCGKAFRSISYFRIHERT HTGEKPYECKCGKAFSRISYFRIHERT HTGEKPYECKECGKAFSRISYFRIHERT HTGEKPYECKECGKAFSRISYFRIHERT HTGEKPYECKECGK CGKAFSCSSYRIHKRTHTGEKPYECKECGK AFIYFTSFQGIMRMHTGEKPYKCKECGKAFSVTS LKSPMRHTIRINYELCE*OLGGKAFSVTS LKSPMRHAQSDRKLY/KCEK*EKVFNSNRCF QSCENSH*REKSCQCK*YKKRDTR*FFMYSQ PHNHVSVSNOPYBCCHEC*OLGGKAFSVTS LKSPMRNAQSDRKLY/KCEK*EKVFNSNRCF QSCENSH*REKSCQCK*YKKRDTR*FFMYSQ PHNHVSVSNOPYBCHEC*OLGGKAFSVTS LKXPMRNAQSDRKLY/KCEK*EKVFNSNRCF QSCENSH*REKSCQCK*YKKRDTR*FFMYSQ PHNHVSVSNOPYBCHCC*OLGKAFSVTS LKXPMRNAQSDRKLY/KCEK*EKVFNSNRCF QSCENSH*REKSCQCK*YKKRDTR*FFMYSQ PHNHVSVSNOPYBCHCC*OLGKAFSVTS LKXPMRNAQSDRKLY/KCEK*EKVFNSNRCF QSCENSH*REKSCQCK*YKKRDTR*FFMYSNLAQHEERERAQKKLEKV MEFEGLKSULQNNL*FSAASMRFOKVFF 443 1793 A 3582 3335 1909 HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLLAQHEERERAQKKLEKV MEFEGLKDEKKLRRSAHARKETEFLRKTT RLGLEDFESLKVIGRGAFGEWRLVQKKDTGH VYAMKLRADML EKEQVGHRAAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEETGFYJAETVLAIDSHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAARTEFYRNLHISLPSDFTFONMNSKRKAE TWKRNRRQLAFSTVGTDSHAPSTVLAIDSHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAARTEFYRNLHISLPSDFTFONMNSKRKAE TWKRNRRQLAFSTVGTDYIAPEVFMQTGYN KLCDWWSLGVIMYSMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WERRIGAPGVEEIKSNFGFGVDWEHREPPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHTE TDYKNKDWFFNTTYKRFGGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGEKRFAVSFLQQLERYDEARQVILEFD GGSRGKGEHFPVEQELKFFAKVVLPLIDQYFK NIRLYYSLAASRPLCSGGHASNKERMVTSL FCKLGVLVBHRISLFGDDATSINCLHILGQDF LDARTTWKTGLESVKSALRAFLDNAAEDLE KTMENLKQGGTHTRNOPKGYTQIINYTTVA LLPMLSSLFEHIGQHQFEGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	İ	i					
FIFPSFLRVHER.HTGEKPYKCK.QCGK.AFR.WS TSIQHERIHTGEKPYKCKECGKSFSARPAFRV HVRVHTGEKPYKCKECGKAFSRISYPRIHERT HTGEKPYECKKCGKTFNYPLDLKIHKRNHTG ERYPECKECACKTEISHRHMITHGDGPY KCRDCGKVFIFPSALRTHERTHTGEKPYECKC CGKAFSCSSYRIHKRTHTGEKPYECKCGCKAFS AFYPTSPGGHIMRMITGERYPYCKECGKAFS LHSSFRRHTRIENYEKPLEC*Q\CGKAFSVSTS LKKPMRNAQSDRLY/KCGSPRLYVT*NISDNRNL VAVVTP*CSTLFKCLWWCKRAALSVV*JVQ DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL ANTVKPHLY 443 1793 A 3578 287 114 DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK LSIKKSVLQNNL*FSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLLANAMMTGSTPCSSMSNITIKERVT MTKVTLENYSNLAQHEEREMRQKKLEKV MEEEGLKDEEKRLRRSAHARKETEFLRLRRT RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKLRKADMLEKFQVGHRAAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEEETOFYJAETVANDSHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTTGNMNSKRKAE TWKNRNRQLAFSTVGTFDYJAPEVFMOTOYN KLCDWSSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHRERPA AISIEKSIDDTSNFDEFPESDILKPTVATSNHIPE TDYKNKDWFINTYTKRFEGLTARGAHPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGGRGGGEHFPVEGEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVBHRISLFGNDATSINCLHILLGGD LDARTYMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGGTTHTRNQPKGYTQINYTTVA LLPMLSSLFEHIGQHQFEGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERGRSALGECLAAFAGA	[
TSIQHERIHTGEKPYKCKECGKSFSARPAFRY HVRVHTGEKPYECKECGKAFSRISYFRHERT HTGEKPYECKECGKTNYPLDLKIHKRNHTG EKPYECKECAKTRISLENFRHMITHTGDGPY KCRDCGKVFIFPSALRTHERTHTGEKPYECKECGKAF KCRDCGKVFIFPSALRTHERTHTGEKPYECKECGKAF SLHSFRRHTRIHNYEKPLEC-QCGKAFSVSTS LKKPMRNAQSDRKLYKCECFAFKCKECGKAF SLHSFRRRHTRIHNYEKPLEC-QCGKAFSVSTS LKKPMRNAQSDRKLYKCECFAFKCKECGKAF SLHSFRRRHTRIHNYEKPLEC-QCGKAFSVSTS LKKPMRNAQSDRKLYKCECFAFKCKECGKAF SLHSFRRRHTRIHNYEKPLEC-QCGKAFSVSTS LKKPMRNAQSDRKLYKCECFAFKCKECGKAF SLHSFRRRHTRIHNYEKPLEC-QCGKAFSVSTS LKKPMRNAQSDRKLYKCECFAFKCKCGGKAFSVSTS LKKPMRNAQSDRKLYKCECFAFKTSSNSTRIL VAVVTP*CSTLFKCLWCWCKRAALSVV*,NVQ DSGRGRWLTPVPBALWAKAGGSRGQEIKTIL ANTVKPHLY 443 1793 A 3578 287 114 DFYERKFEGFIEHKQIVNKWRDLLCSWKRK LSILKKSVLQNNL*PSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAAMANTGSTPCSSMSNHTKERVT MTKVTLENFYSNLAQHEREMRQKKLEX WEEEGLKDEEKRLRRSAHARKETEFLRKRT RLGLEDFESLKVIGRGAFGEVLVQKKDTGH VYAMKLIRKADMLEKEQVGHRAAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEETQFYJAETVVLAIDSHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLTCGLK KAHRTEFYRNLHISLFSDFTFONMNSKRKAE TVKRNRRQLAFSTVGTFDYJAPEVYMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLLRFCCE WERRIGAPGVEEKISNSFFEGVDWEHRERPA AISIEKSIDDTSNFDEFPESDILKPTVATSNHEE TDYKNKDWFNYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGEKRFAYSFLQQLIRYVDEAHQYILEFD DSKRYMKGGLESVEKSFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSINCLHILGGE KTMENLKQGGFTHTRNQPKGYTQINYTTVA LLPMLSSLFEHIGQHOFEGEDLILEDVQVSCYRI LDARTTWKKTGLESVEKSALRAFLDNAAEDLE KTMENLKQGGFTHTRNQPKGYTQINYTTVA LLPMLSSLFEHIGQHQFFEGDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA		1					
HVRVHTGEKPYKCKEGGKAFSRISYFRIHERT HTGEKPYECKCCKCKTFINYPLDLKHKRNHTG EKPYECKECAKTISLENFRRHMITHTGDGPY KCRDCGKVFIFFSALRTHERTHTGEKPYECKC CGKAFSCSSYRIKRTHTGEKPYECKCGKAFS LHSSFRURHTRIHNYEKPLEC*QCGKAFSVSTS LKSFRURHTRIHNYEKPLEC*QCGKAFSVSTS LKSFRURHTRIHNYEKPLEC*QCGKAFSVSTS LKSFRURHTRIHNYEKPLEC*QCGKAFSVSTS LKKPRMNAGSNEK/VFCK*VEKVENSNCF QSCENSH*REKSCQCK*YRKRDTR*FMYSQV PHNHVSVSNGPYRCCSPRILYNT*NISINRNL VAVVTPCSTLFKCLWCWCKRAALSVV*7/VQ DSGRGRWLTPVTPALWEAKAGGSRGQEIKTIL ANTVKPHLY 443 1793 A 3578 287 114 DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK LSIIKKSVLQNNL*FSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEEGLKDEEKRLRSAHARKETEFLRLKRT RLGLEDFESLKVIGGAFAGFEVRLVQKKDTGH VYAMKLRKADMLEKEQVGHRAERDILVEA DSLWVVKMFYSFQDKLINLYLMFEFLPGGDM MTLLMKKDTLTEETQFYIAETVLQKKDTGH VYAMKLRKADMLEKEQVGHRAERDILVEA DSLWVVKMFYSFQDKLINLYLMFFLPGGDM MTLLMKKDTLTEETGFYTAETVLAIDSHQL GFHRDIKPDNLLLDSKGFVKLSDFGLCTGLK KAHRTEFYRNLNHSLEFDFTFQNMNSKRKAE TWKRNRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPFCSSTPQETY KKVVMWKETLTPPEVPISEKAKDLILRFCCE WEHRIGAFGVEEKSNSFFEGVDWEHRERPA AISIEKSIDDTSNTDEFPESDILKPTVATSNHTE TDYKNKDWFINTYTYKREGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDĒAHQYILEFD GGSRGGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRIRISLFGNDAATSINCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSTYVERQRSALGECLAAFAGA			1		,		
HTGEKPYECKEGKTENYPLDI.KIHKRNHTG EKPYECKECAKTFISLENFRRHMITHTGDGPY KCRDCGKVFIFPSALRTHERTHTGEKPYECKE CGKAFSCSSYRIHKRTHTGEKPYECKECGK AFIYPTSFQGHMRMHTGEKPYECKECGKAFS LHSSFRURTRIHNYEKPLEC*Q\CGKAFSVSTS LKKFMRNAQSDRKLY/KCEK*EKVFNSNRCF QSCENSH*REKSCQC*YRKRDTR*FMYSQV PH>HVSVSNGPYRCGSPRLYNT*NISNRNL VAVVTPYCSTLFKCLWCWCKRAALSVV*/IVQ DSGRGWLTPVPALWEAKAGGSRGQEIKTIL ANTVKPHLY 443 1793 A 3578 287 114 DFYERKFEGFEGHKQIVNKWRDLLCSWKRK LSILKSVLQNNL-PSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLAQHEERMQKKLEV MEEEGLKDEEKRERSAHARKETEFLRLKRT RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKLRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSOLKINLYLINEFLFGGDM MTLLMKKDTLTEETQFYJAETVLAIDSHQL GFHRDIKPDNLLLDSKGHKLSVFGTPQTY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHRERPA AISIEKSIDDTSNFDEFPESDILKPTVATSNHTE TDYKKNDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKCGHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMYTSL FCKLGVLVRHRISLFONDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKGQGFTHTRNQPKGVTQINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSTYVERQRSALGECLAAFAGA	1						
EKPYECKECAKTFISLENFRRHMITHTGGGPY KCRDCGKVFIFPSALRTHERTHTGEKPYECKC CGKAFSCSSVIRHKRTHTGEKPYECKECGK AFIYPTSFQGHMRMHTGEKPYECKECGK AFIYPTSFQGHMRMHTGEKPYKCKECGKAFS LHSSFRAHRTHNYEKPLEC*QCGKAFSVSTS LKKFMRNAQSDRKLY/KCEK*PKKVFNSNRCF QSCENSH*REKSCQCK*YRKRDTR*FMYSQV PHNHVSVSNGFYRCGSPIRLYNT*NISINRNL VAVVTPCSTLFKCLWCWCKRAALSVV*7IVQ DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL ANTVKPHLY 443 1793 A 3578 287 114 DFYERKFEQFEGHKQIVNKWRDLLCSWKRK LSIIKKSVLQNNL*FSSASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAMAMTGSTPSSMSNHTKERVT MTKVYLENFYSNLIAQHEEREMRQKKLEKV MEEEGLKDEEKLRRSAHARKETEFLRLKRT RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHRAERDIL VEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEEETGYTVAETVLAIDSHQL GFHRDIKPDNLLIDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRAE TWKRNRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWSLGVINYEMLIGYPFFCSSTPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEKSNSFFEGVDWEHRERPA AISIEKSIDDTSNFDEFPESDILKPTVATSNHFE TDYKNKDWVFNYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	İ					,	
KCRDCGKVFIFPSALRTHERTHTGEKPYECKC CGKAFSCSSYRIHKRTHTGEKPYECKCCGK AFIYPTSFQGHMRMHTGEKPYKCKECGKAFS LHSSFRKHTRIHNYEKPLEC*Q\CGKAFSVSTS LKKPMRNAQSDRKLY/KCEK*EKVFNSNRCF QSCENSH*REKSCQCK*YKRDTR*FMYSQV PHNHVSVSNGPYR/CGSPRLYNT*NISINRNL VAVVTP*CSTLFKCLWCWCKRAALSVV*/IVQ DSGRGRWLITPVIPALWEAKAGGSRGGEIKTIL ANTVKPHLY 443 1793 A 3578 287 114 DFYERKFEGFIEGHKQIVNKWRDLLCSWKRK LSIIKKSVLQNNL*FSAASMFQKVFF 444 1794 A 3582 3335 1909 HLFFSLLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLLAQHEEREMRQKKLEKV MEEEGLKDEEKRLRRSAHARKETEFLRLKRT RLGLEDFFSLKVIGRGAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEETQFYIAETVLAIDSHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLHSLPSDFTFQMNNSKRKAE TWKRNRRQLAFSTVGTTDYIAPEVFMQTGYN KLCDWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRTCCE WEHRIGAFQVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHTE TDYKNKDWVFNYTYKREGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMYTSL FCKLGVLVHRIBLSLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGGTHTIRNOPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA		Ī					
CGKAFSCSSYRIHKRTHTGEKPYECKECGK AFIYPTSFQGHMRMHTGEKPYKCKECGKAFS LHSSFRRHTRIHNYEKPLEC*Q\CGKAFSVSTS LKKPMRNAQSDRKLY/KCEK*EKVPNSNRCF QSCENSH*REKSCQCK*YRKRDTR*FMYSQV PHNHVSVSNGPYR/CGSPIRLYNT*NISINRNL VAVVTP*CSTLFKCLWCWCKRAALSVV*/IVQ DSGRGWLTPVIPALWEAKAGGSRGQEIKTIL ANTVXPHLY 443 1793 A 3578 287 114 DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK LSILKKSVLQNNL*PSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAAMTGSTPCSSMSNHTIKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEEGLKDEEKRLRSAHARKETEFLRLKRT RLGLEDFSILKVIGRGAFGEVRLVQKKDTGH VYAMKLRADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEETQFYIAETVLAIDSHQL GFHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQMMSKRKAE TWKRNRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHTE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIECRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRNYFISAASRPLCSGGHASNKEKEMYTSL FCKLGVLVWHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGGFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLLLEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA		1					
AFIYPTSFQGIMRMHTIGEKPYKCKECGKAFS LHSSFRRHTRIHNYERJEC*Q(CGKAFSVSTS LKKPMRNAQSDRKLY/KCEK*EKVPNSNRCF QSCENSH*REKSCQCK*YRKRDTR*FMYSQV PHNHYSVSNGPYB/CGSBRILYNT*NISINRNL VAVVTP*CSTLFKCLSWCKKAALSVV*JVQ DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL ANTVKPHLY 443 1793 A 3578 287 114 DFYERKFEGFIEGHKQIVNKWRDLLCSWKRK LSIIKKSVLQNNL*FSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEGGLKDEEKRRSAHARKETEFLR.KRT RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKLIRKADMLEKEQVCHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEETGPYIAETVLAIDSHQL GFHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRQLAFSTVGTPDYJAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTPPEVPISEKAKDLIRFCCE WEHRIGAFQVEEIKSNSFFEGVDWEHRERPA AISIEKSIDDTSNFDEFPESDILKPTVATSNHTE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPVEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNEKEMTYSL FCKLGVLVWHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGGFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLLLEDVQVSCYQL LLPMLSSLFEHIGQHQFGEDLLLEDVQVSCYQL LLSSLFSHIGQHGFGEDLLLEDVQVSCYQL LLSSLFSHIGQHGFGEDLLEDVQVSCYQL LTSLYALGTSKSIYVERQRSALGECLAAFAGA		1			:		
LHSSFRIRHTRIHNYEKPLEC*Q(CGKAFSVSTS LKKPMRNAQSDRKLY/KCEK*EKVPNSNRCF QSCENSH*REKSCQCK*YRKRDTR*FMYSQV PHNHVSVSNGPYR/GSPRLYNT*NISINRNL VAVVTP*CSTLECHWCWCKAALSVV*/IVQ DSGRGRWLTPVPALWEAKAGGSRGQEIKTIL ANTVKPHLY 443 1793 A 3578 287 114 DFYERKFEQFIEGHKQIVNKWRDLLCSWKR LSIIKKSVLQNNL*PSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEGLKDEEKLRRSAHARKETEFLRLKT RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEETQFYIAETVLAIDSHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKNNRQLAFTVGTPDYLAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISEIKSIDDTSNFDEFPESDILKPTVATSNHTE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKFFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYTLSAASRPLCGGHASNKEKEMYTSL FCKLGYLVRIRISLFGNDATSINVCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHITNQPKGYTQIINYTTTVA LLPMLSSLFEHIGQHOFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA		1	•				
LKKPMRNAQSDRKLY/KCEK*EKVFNSNRCF QSCENSH*REKSCQCK*YRKDTR*FMYSQV PHNHVSVSNGPYR/CGSPIRLYNT*NSINRNL VAVVTP*CSTLFKCLWCWCKRAALSVV*/IVQ DSGRGRWLTPVPALWEAKAGGSRGQEIKTIL ANTVKPHLY 443 1793 A 3578 287 114 DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK LSIIKKSVLQNNL*PSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAAMAMTGSTPCSSNSHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEGLKDEEKRLRRSAHARKETEFLRLKRT RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEETQFYIAETVLAIDSHQL GFHRDIKPDNLLLDSKGHVKLSPGLCTGLK KAHRTEFYRILNSLPSDFTFONMNSKRKAE TWKRNRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGGHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRIRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHIRNQPKGVTQIINYTTTVA LLPMLSSLFEHIGQHOFGEDLLEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	,				•		AFIYPTSFQGHMRMHTGEKPYKCKECGKAFS
QSCENSH*REKSCQCK*YRKRDTR*FMYSQV PHNHVSVSNGPYR/CGSPRLYNT*NISINRNL VAVVTPPSTLFKCLWCWCKRAALSVV*/IVQ DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL ANTVKPHLY 443 1793 A 3578 287 114 DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK LSIIKKSVLQNNL*FSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEEGLKDEEKRLRSAHARKETEFILKRT RIGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEETQFYIAETVLAIDSHQL GFIHRDIKRPDNLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRQLAFSTVGTPDYIAPEVFMOTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEKSIDDTSNFDEFPESDILKPTVATSNHTE TDYKNKDWVFNYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLYNHRISLFGNDATSIVNCLHILQQT LDARTVMKTGLESVKSALRAFLDNAÆDLE KTMENLKQQGFHTRNORKGVTQINNYTTVA LLPMLSSLFFHIGGHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA		1					LHSSFR\RHTRIHNYEKPLEC*Q\CGKAFSVSTS
PHNHVSVSNGPYR/CGSPIRLYNT*NISINRNIL VAVVTP*CSTLFR/CLWCWCKRAALSVV*/IVQ DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL ANTVKPHLY 443 1793 A 3578 287 114 DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK LSIIKKSVLQNNL*FSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEEGLKDEEKRLRSAHARKETEFLRLKRT RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEETQFYIAETVLAIDSHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISEIKSIDDTSNFDEFPESDILKPTVATSNHTIE TDYKNKDWVFINTTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPVEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILQQT LDARTVMKTGLESVKSALRAFLDNAÆDLE KTMENLKQQGFTHTRNQPKGGVTQINYTTVA LLPMLSSLFFHIGGHQFGEDLILEDQVGVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA							
VAVVTP*CSTLFKCLWCWCKRAALSVV*/IVQ DSGRGRWLTPVPALWEAKAGGSRGQEIKTIL ANTVKPHLY 443 1793 A 3578 287 114 DFYERKFEOFIEGHKQIVNKWRDLLCSWKRK LSIIKKSVLQNNL*FSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEGGLKDEEKRLRSSAHARKETEFLRLKRT RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEEETOFYIAETVLAIDSHQL GFHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEIKSNSFFEGVDWEHIRERPA AISIEKSIDDTSNTDEFPESDILKPTVATSNHTEE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQINYTTYA LLPPILSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	ł				•		
DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL ANTYKPHLY 443 1793 A 3578 287 114 DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK LSIKKSVLQNNL*PSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEEGLKDEEKRLRRSAHARKETEFLRLKRT RLGLEDFESLKVIGRGAFGEVRL*VQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGDDM MTLLMKKDTLTEEETOFYIAETVLAIDSHQL GFHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEKSNSFFEGVDWEHIRERPA AISIEKSIDDTSNFDEFPESDILKPTVATSNHIPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHIPPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGGFTHTRNQPKGVTQIINYTTYA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	1	1			•		
443 1793 A 3578 287 114 DFYERKFEQFEGHKQIVNKWRDLLCSWKRK LSIKKSVLQNNL*FSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEGGLKDEEKRLRRSAHARKETEFLRLKRT RLGLEDFESLKVIGRAGFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEETQFYIAETVLAIDSHQL GFHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPBVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEKSNSFFEGVDWEHIRERPA AISIEKSIDDTSNFDEFFESDILKPTVATSNHIPE TDYKNKDWVFNYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKGGFTHTRNOPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA		l :					VAVVTP*CSTLFKCLWCWCKRAALSVV*/IVQ
443 1793 A 3578 287 114 DFYERKFEQFIEGHKQIVNKWRDLLCSWKRK LSIIKKSVLQNNL*FSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEGGLKDEEKRLRRSAHARKETEFLRLKRT RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEETQFYIAETVLAIDSIHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHTE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPVEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGGFTHTRNQPKGVTQIINYTTVA LLFMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA		1					DSGRGRWLTPVIPALWEAKAGGSRGQEIKTIL
LSIKKSVLQNNL*FSAASMRFQKVFF 444 1794 A 3582 3335 1909 HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEGLKDEEKRLRRSAHARKETEFLRKRT RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTTEETQFYIAETVLAIDSIHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTTQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEKSNSFFEGVDWEHIRERPA AISIEKSIDDTSNFDEFPESDILKPTVATSNHTPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQQGFTHTRNQPKGVTQIINYTTVA LLPPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA		<u> -</u>					
444 1794 A 3582 3335 1909 HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEEGLKDEEKRLRRSAHARKETEFLRLKRT RLGLEDFESLK VIGRGAFGEVRL VQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEETQFYIAETVLAIDSIHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHTPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	443	1793	Α	3578	287	114	
MTKVTLENFYSNLIAQHEEREMRQKKLEKV MEEEGLKDEEKRLRRSAHARKETEFLRLKRT RLGLEDFESLKVIGRGAFGEVRL VQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEEETQFYIAETVLAIDSIHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHTPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA		· ·					LSIIKKSVLQNNL*FSAASMRFQKVFF
MEEEGLKDEEKRLRRSAHARKETEFLRLKRT RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLINLYLIMEFLPGGDM MTLLMKKDTLTEETQFYIAETVLAIDSIHQL GFHRDIKPDNILLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHIPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	444	1794	Α	3582	3335	1909	HLFFSLFLAAMAMTGSTPCSSMSNHTKERVT
RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEETQFYIAETVLAIDSHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA			1				MTKVTLENFYSNLIAQHEEREMRQKKLEKV
VYAMKILRKADMLEKEQVGHIRAERDILVEA DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEEETQFYIAETVLAIDSIHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTPPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHIPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA							MEEEGLKDEEKRLRRSAHARKETEFLRLKRT
DSLWVVKMFYSFQDKLNLYLIMEFLPGGDM MTLLMKKDTLTEETQFYIAETVLAIDSIHQL GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHTE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQINYTTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA					ļ		RLGLEDFESLKVIGRGAFGEVRLVQKKDTGH
MTLLMKKDTLTEBETQFYIAETVLAIDSIHQL GFIHRDIKPDNILLDSKGHVKLSDFGLCTGLK KAHRTEFYRNI.NHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHTE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQINYTTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA						ļ	VYAMKILRKADMLEKEQVGHIRAERDILVEA
MTLLMKKDTLTEBETQFYIAETVLAIDSIHQL GFIHRDIKPDNILLDSKGHVKLSDFGLCTGLK KAHRTEFYRNI.NHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHTE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQINYTTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA					l		
GFIHRDIKPDNILLDSKGHVKLSDFGLCTGLK KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHIPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA						İ	
TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	1			ľ	ļ		GFIHRDIKPDNLLLDSKGHVKLSDFGLCTGLK
KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	1		i		i		KAHRTEFYRNLNHSLPSDFTFQNMNSKRKAE
KLCDWWSLGVIMYEMLIGYPPFCSETPQETY KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA							TWKRNRRQLAFSTVGTPDYIAPEVFMQTGYN
KKVMNWKETLTFPPEVPISEKAKDLILRFCCE WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHIPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA		ļ·					
WEHRIGAPGVEEIKSNSFFEGVDWEHIRERPA AISIEIKSIDDTSNFDEFPESDILKPTVATSNHPE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 I 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA		<u> </u>	ļ	j			
AISIEIKSIDDTSNFDEFPESDILKPTVATSNHTE TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	1		i	ŀ			
TDYKNKDWVFINYTYKRFEGLTARGAIPSYM KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASIVNCLHILIGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTIVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA		1	1	ļ	l		
KAAK 445 1795 A 3584 1 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA		1 1					
445 1795 A 3584 I 6169 RTRGIEKRFAYSFLQQLIRYVDEAHQYILEFD GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	1	, 1			1		
GGSRGKGEHFPYEQEIKFFAKVVLPLIDQYFK NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	445	1795	A	3584	1	6169	
NHRLYFLSAASRPLCSGGHASNKEKEMVTSL FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	1		1		1	- /	
FCKLGVLVRHRISLFGNDATSIVNCLHILGQT LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	1				1		
LDARTVMKTGLESVKSALRAFLDNAAEDLE KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	!			1	l		
KTMENLKQGQFTHTRNQPKGVTQIINYTTVA LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	İ						
LLPMLSSLFEHIGQHQFGEDLILEDVQVSCYRI LTSLYALGTSKSIYVERQRSALGECLAAFAGA	İ	f [l	Ī	1	
LTSLYALGTSKSIYVERQRSALGECLÄAFAGA				i	ŀ		TI DIA COI ECUICOUOPCEDI A CONOCCEDI
				l	l		
, i i i i prvahlelmijkmnitnitnikankedaalcid	1				j	İ	
	l			l	İ	ľ	
TNVEDVCPNIPSLEKLMEEIVELAESGIRYTQ	L	<u> </u>			l		INVEDVCPNIPSLEKLMEEIVELAESGIRYTQ

						PRLVSNSWTQAILLPRPPKMLGLQV
			}		ŀ	DTVLLCCSGWSAVAPSRLTAALFS*AQAVCL SLPRSWDYRRW/PPHPANFCIFCRDE/SLA/ML
446	1796	A	3592	1	355	AGLELLNSDDPPALASQSAGITGVTRTPSLFF*
			- 1		İ	FLMYLINKDETEHTGQESYVWKMYQERCWE FFPAGDCFRKQYEDQLN
	-)		ļ	FICGIGNDYFDTVPHGFETHTLQEHNLANYLF
	1				1	VILLAIIQGLIIDAFGELRDQQEQVKEDMETKC
			}			YNKSEDGDTPDMKCDDMLTCYMFHMYVGV RAGGGIGDEIEDPAGDEYEIYRIIFDITFFFFVI
1	1		- 1		1	NGKQLVLTVGLLAVVVYLYTVVAFNFFRKF
1		Ì				VKYQMWKLGVVFTDNSFLYLAWYMTMSVL GHY\NNFFFAAHLLDIAMGFKTLRTILSSVTH
- 1	İ		1	İ	ł	MDKAALDFSDAREKKKPKKDSSLSAVLNSID
l	. 1	•	- 1	•		LEFDGLYITEQPSEDDIKGQWDRLVINTQSFP NNYWDKFVKRKVMDKYGEFYGRDRISELLG
	- 1		1			PILHTVISFFCIIGYYCLKVPLVIFKREKEVARK
			ļ	1		AFAINFILLFYKVSTSSVVEGKELPTRSSSENA KVTSLDSSSHRIIAVHYVLEESSGYMEPTVRIL
	1]			FWKKIIAYQQKLLNYFARNFYNMRMLALFV
1	[-			KEEKAKEDKGKQKLRQLHTHRYGEPEVPESA
ſ	ĺ	1	- (İ		EGGQYKLIPHNPNAGLSDLMSNPVPMPEVQE KFQEQKAKEEEKEEKEETKSEPEKAEGEDGE
	ļ	l				LEAALPSEDLTDLKELTEESDLLSDIFGLDLKR
				İ		FMTLLHFVASVFRGFFRIICSLLLGGSLVEGA KKIKVAELLANMPDPTQDEVRGDGEEGERKP
	į	1	,			LKSLKKQMKKVKKMTVKDMVTAFFSSYWSI
ļ]					TIFEMQLAAQISESDLNERSANKEESEKERPEE QGPRMAFFSILTVRSALFALRYNILTLMRMLS
			ĺ			VKESKRQFIFDVVNEGGEKEKMELFVNFCED
ſ	j	j	- [VLLTNLSEHMPNDTRLQTFLELAESVLNYFQP FLGRIEIMGSAKRIERVYFEISESSRTQWEKPQ
	ļ	1	{			SCAETDENETLDYEEFVKRFHEPAKDIGFNVA
			1			PDGKGVIFKRDFHKAMESHKHYTQSETEFLL
1	ł			,	i	KDMVVMLLSMLEGNVVNGTIGKQMVDMLV ESSNNVEMILKFFDMFLKLKDLTSSDTFKEYD
	1					VGFLHVFAHMQMKLSQDSSQIELLKELMDLQ
j						ESISDFYWYYSGKDVIDEQGQRNFSKAIQVA KQVFNTLTEYIQGPCTGNQQSLAHSRLWDAV
						NSDFQNYLRTQTGNNTTVNIIISTVDYLLRVO
]					FQSLAGLMQSCSVLDLNAFERQNKAEGLGM VTEEGSGEKVLQDDEFTCDLFRFLQLLCEGH
				İ		LKLGIAILNGGNSTVQQKMLDYLKEKKDVGF
	ĺ			ļ		ARLHDRGAAEMVLQTISASKGETGPMVAAT
						LILLFSRTALTEKCKLEEDFLYMAYADIMAKS CHDEEDDDGEEEVKSFEEKEMEKQKLLYQQ
Ì						YFEDKLIEDLAKPGAEPPEEDEGTKRVDPLHQ
						CLVEHPQRSKKAVWHKLLSKQRKRAVVACF RMAPLYNLPRHRAVNLFLQGYEKSWIETEEH
	ļ					DPEKTVERVLDIANVLFHLEQKSKRVGRRHY
						NICAPGDQELIALAKNRFSLKDTEDEVRDIIRS NIHLQGKLEDPAIRWQMALYKDLPNRTDDTS
						ERKKMKRKGDRYSMQTSLIVAALKRLLPIGL
						QNFVVQNEINNMSFLITDTKSKMSKAAVSDO
						LILDEFTTLARDLYAFYPLLIRFGDYNRAKWL KEPNPEAEELFRMVAEVFIYWSKSHNFKREE
			l i		}	EKLKKKAATVVSEEDHLKAEARGDMSEAEL
						AEMCCTALNSEHMNTLLGNILKIIYNNLGIDE GAWMKRLAVFSQPIINKVKPQLLKTHFLPLM
						MPHVMEVILPMLCSYMSRWWEHGPENNPER
				sequence		nucleotide insertion
	ļ.			residue of peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
:				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
uence	201100	[914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
eotide seq-	seq- uence		USSN 09/496	location	corresponding to last amino	I=Isolcucine, K=Lysine, L=Lcucine,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.

Sey III of peptide enable peginning of the peginning of the peginning of the peginning of the peginning of the peginning of the peginning of the peginning of the peginning of the peginning of the peginning of the peginning of the peginning of the peginning of the peginning of p	CCC ID	L CEO ID	17.4	l cco	T 10 - 11 - 1	130 10 1 1 1	14 : 6 6
mucloide cofide sequence wence where the contemporary of the conte	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence USN 09/496 19/4 19 19 19 19 19 19 19 1		1	1100				
1797 A 3598 1202 1070 1797 Linux mino acid residue of pepide residue of pepide sequence Pepide sequence		1 * *	Í	1			
1914 ng to first amino acid residue of peptide sequence 17-Threosine, V-Valine, W-Pryptophan, Y-Tyrosine, X-Unknown, "-Stop codon, Peptide sequence 17-Threosine, V-Valine, W-Pryptophan, Y-Tyrosine, X-Unknown, "-Stop codon, Peptide sequence 17-Threosine, V-Valine, W-W-Yrytophan, Y-Tyrosine, X-Unknown, "-Stop codon, Peptide in the International Color of the Inte	1		1				M-Methiania N. Assessing D. D. Line
amino acid residue of peptide residue of peptide sequence peptide sequence peptide sequence per control of the period sequence per control of the period sequence per control of the period sequence per control of the period sequence per control of the period sequence per control of the period sequence per control of the period of the per		nettee	l				M=Methorine, N=Asparagine, P=Profine,
Presidue of peptide sequence	ucnce	İ	l	314			
Pepside sequence	ļ	ļ					
1797 A 3598 1202 1070 LFVGGGPCEGASGFAFGFAPRAPRYGVDAEV GR*V*GAAASQGA/GSLRRPTGGBIFGAWL QWGAAAASQGA/ASLRRPTGGBIFGAWL QWGAAAASQGA/ASLRPTGGBIFGAWL QWGAAAASQGA/ASLRPTGGBIFGAWL QWGAAAASQGA/ASLRPTGGBIFGAWL QWGAAAASQGA/ASLRPTGGBIFGAWL QWGAAAASQGA/ASLRPTGGBIFGAWL QWGAAAASQGA/ASLRPTGGBIFGAWL QWGAAAACGAGAM*YARARGBRAGHLCYI NGTLALGILEP*4*WGWGEWRENG GPSRSAPGLASGARALQGAYSKAPHLCYI NGTLALGILEP*4*WGWGEWRENG GVEPLFVNNVKVERHGGRWVVLAAVLIG LLI*VILGIGFILWHI,QYRDVRYQKYFNGYM RTRINGFYDAYAFNANGE EVALSKEYKDALK LYSGVPFLGFTHESA YTAFSGGSYLAYYWS FRRENGGGFKDFGAGLKYYSISHEKYVALKE LYSGVPFLGFTHESA YTAFSGGSYLAYYWS FRRENGGFKDFGAAKYGTDNSCSGGLAR GVELMFTTFGFPDFFYFAHAACQWALRGD ADSVLSLTFRSFDLASCDERGRHLVTYYNTUL LITLITTTERRHPGFEATFGLFRNSCGGRL RKAQGTTNSPYYGGHYPHDICTVNIELPW LITLITTTERRHPGFEATFGLFRNSCGGRL RKAQGTTNSPYYGGHYPHDICTVNIELPW LITLITTTERRHPGFEATFGLFRNSCGGRL RKYGGRSGFYYTSNSHYTHTSPSYDYFHTHSPHDCTVNIELPW GENYGGRSGFYYTSNSHYTHTSPSYDYFHTHSPHDCTVNIELPW GENYGGRSGFYYTSNSHYTHTSPSYDYFHTHSPHDCTVNIELPW GENYGGRSGFYYTSNSHYTHTSPSYDYFHTHSPHDCTVNIELPW GENYGGRSGFYYTSNSHYTHTSPSYDGFTYTHTSPHDCTVNIELPW GENYGGRSGFYYTSNSHYTHTSPHDCTVNIELPW GENYGGRSGFYYTSNSHYTHTSPHDCTVNIELPW GENYGGRSSFYYTSNSHYTHTSPHDCTVNIELPW GENYGGRSSFYYTSNSHYTHTSPHDCTVNIELPW GENYGGRSSFYTTSNSHYTHTSPHDCTVNIELPW GENYGGRSSFYTTSNSHYTHTSPHDCTVNIELPW GENYGGRSSFYTTSNSHYTHTSPHDCTVNIELPW GENYGGRSSFYTTSNSHYTHTSPHDCTVNIELPW GENYGGRSSFYTTSNSHYTHTSPHDCTVNIELPW GENYGGRSSFYTTSNSHYTHTSPHDCTVNIELPW GENYGGRSSFYTTSNSHYTHTSPHDCTVNIELPW GENYGGRSSFYTTSNSHYTHTSPHDCTVNIELPW GENYGGRSSFYTTSNSHYTHTSPHDCTVNIELPW GENYGGRSSFYTTSNSHYTHTSPHDCTVNIELPW GENYGGRSSFYTTSNSHYTHTSPHDCTVNIELPW GENYGGRSSFYTTSNSHYTHTSPHDCH GENYGGRSFYTSNSHYTHTSPHDCH GENYGGRSFYTTSNSHYTHTSPHDCH GENYGGRSFYTSTTVLTGAA ASITSVDPAGHHIDATAGAGHGSSFYTTSNSHYTHTSPHDCH GENYGGRSFYTSTTVLTGAA ASITSVDPAGHHIDATAGAGHGSSFYTSTTVLTGAA ASITSVDPAGHHIDATAGAGHGSSFYTSTTVLTGAA ASITSVDPAGHHIDATAGAGHGSSFYTSTTVLTGAA ASITSVDPAGHHIDATAGAGHGSSFYTSTTVLTGAA ASITSVDPAGHGSTTSSTAGEE GENYGGRAGGACHAWA	ĺ	1	(1		sequence	
1997 A 3598 1202 1070	ţ	İ	ľ	1		1	
GR*V*GAAASQGA/GSLPRRTOPGHPGAWL QVWGAAAVCAGPAM/AVRAKRGPRAGEP NSPWRSGVLAAURAVGAGPPWP*PPGCS*PARE SPSSSAPGLASGPAAPLLQCVHSAGPLLCYI MSTLALGLKP*AWGWGEWRYKG GVELPVNNKKEVEHGPGSWVLAAVLIG LLIVLLGIGFLVWHLQYRDVRQKVPHGYW KITNENFVDAYENNSTEPVSLASKVBDALKL LVSGVPFLCPVHKESAVTAFSEGSVAYYWSE SPIPQRLVEAREWARERVYMLPPARSILKS FVVTSVVAPTDSKTVQRTQDNSCSGGLHAR GVELMRTTIGPGPSYPAHARCQWALRGD ADSVLSLTRSFDLASCDERGRHLVTVYTYTLL SPMEPHALVQLGGTYPSYNLFHSISSQNVL LITLITITEREHPGVEATFFQLPEANSLCSGGRL RKAQGTPSSPYYPGHARCQWALRGD ADSVLSLTRSFDLASCDAGRGHLVTVYYTL SPMEPHALVQLGGTYPSYNLFHSISSQNVL LITLITITITEREHPGVEATFFQLPEANSLCGGRL RKAQGTPSSPYYPGHALVQLGGTYPSNVLFHSDQSYTDT GREATIST GREATI	447	1707		3500		1070	
QVWGAAAVCAGPAM*ARRAGRAGPHP*PFOGCS*ARG PSSRSAPGLASGPAAPHCGCS*ARG PSSRSAPGLASGPAAPHCGS*ARG PSSRSAPGLASGPAAPHCGS*ARG PSSRSAPGLASGPAAPHLOGHSSAGPLLCYI NGTLALGLR***AVGWGEWRPKC GVELPYNNIKXUEWHGPGRWYUAAVLG LLJ*VLLGIGFLVWHLQYRDYRVGKVYNGCYM RITNEN*PUAAYBNSTEPVSLASK*KADALKL LYSGVPFLQPYNNIKXUEWHGPGRWYUAAVLG LLJ*VLLGIGFLVWHLQYRDYRVGKVYNGCYM RITNEN*PUAAYBNSTEPVSLASK*KADALKL LYSGVPFLQPYNNIKXUEWHGPGRWYUALPFARSILKS FVYTSVVAPFIDSKTVGRTQNSCSSGLIAR GVELMR*TITGFDSFYFAHARCQWALRGD ADSVLSLTRSFGLASCDEGRRHLVTYNTTL SPMEPHALVQLGTYPPSYNLTHISISGNVL LITLITITITERRIFIGGETAFTFQLPMSSCGGL RKAQGTRNSPYPOHYPPNIDCTWNIEVPNN QHVKVPRKFFYLLEFGYPAGTCPKDYVENG EKYCGERSGPVVTSNSNKITVRFHSDQSYTDT GRLASTLYSTDSSPCFGQFTCRTGRKELR CDGWADCTDHSDELNCSCDAGHGFTCKNEK CPFLFWVCDSLNDCGDDSDSGGGCSCPAQTF RCSNGKCLSKSQQCNGKDDCDGDGSDSASCP KNVVYTCTKHTYRCLNGLLSKGNPECDGK EDCSDSGBEKDCDCGLRSFTRQARVVGGTD ADGEBYWQVSLBLALGQGHLGCASLSPNWL VSAAHCVIDDRGFRYSDPTQWTAFLGLHDGS QRSAPGVGBRLKRIEJSPFRDHFTPOTDIALL ELEKPASYSSWAPPICLDASHVFPAGKAHW TOWGHTDYGGTGALLGKGERW NOPTTCEN LLPQQTPRMMCVGFLSGGVDSCQGDSGGFL SSVEADGRIPQAGVVSWGDGCAQRNKPGVY TRPLFRTWWESHTOV TOWGHTDYGGTGALLGKGERW NOPTTCEN LLPQQTTPRMMCVGFLSGGVDSCQGDSGGFL SSVEADGRIPQAGVVSWGDGCAQRNKPGVY TRPLFRTWWESHTOV APDILPHFLGEFEDVYIVKNKRPULVCKAV APOLIPFFCNGEWWRQVGWNAQQSATVANPVPG ANPDLLPHFLGEFEDVYIVKNKRVLLVCKAV SSGTTLSGAATNIATVKNPFQGFAKEVSL EQGIVLPCRPPEGIPPAE ### MPPS A 3618 2 613 FYSGSPWRNDGSTERLEARPAGRLPWSSRQ EMTRETSHAAGRGHWSAQQSATVANPVPG ANPDLLPHFLGEFEDDVYIVKNKRVLLVCKAV SSGTTLSGAATNIATVKNPFQGFAKEVSL EQGIVLPCRPPGIPPAE ### MPPS A 3620 1 2676 MPPS GGGMDLTCPFGVSPACGAQASWSIFG ADABEVFGTGGISQQEAAMPHPEDEEPPGG POAAQSSAQQQFTAAPVSCPPTTITVLTGDA TSPEGETDKNLANRVHSPHKRLSIRRHLKVST ASLTSDVAGAGGGMDLTCPFGVSPACGAQASWSIFG ADABEVFGTGGISQQEAAMPHPEDEEPPGG POAAQSSAQQQGFTAAP RPPVEKSKELABECKERDPLQYPTETTKCLA PYTNSSGKMALNSPOPGVSSECDEANX LALLEEAKLVSERPLTRRGRKSSSSPGDSPSA VSPNLSPSASTTSSSSSTLTYPTPEGDEADVS SPFHGEPNVPKGLADRKQNGCTAFAP RPPVEKSKELABECKERDPLQYPTETTKGLA PYTNSSGKMALNSPOPGVSSECHAME APGLKDPGAAGQGGFTAGEP MCPGAGGAGAGCPFTLKSKLEHLMNN QNLVGLKLPDLSEAARGEKGLGLFSLSTALEEB ENSGLDVMPINSDVLL	447	1/9/	^	3398	1202	1070	
NSPWRSGVLAARAVGAGPWPPPPGCSYARG PSSRSAPGLASGPAAPLLQVHSSAGPLLCYI NGTLALGLRP*AWGWGEWRKG FSRGGGGERDFGGAGLYNGSEHEKVNGLEE GVEFLPVNNVKKVEKHGPGRWVVLAAVLIG LLIVLLGIGFLVWHLQYRDVRVQKVFRGYW RTNEWFVDAYENNSTEPVSLASKVKDALK LVSGVPFLGPYHKESAVTAFSGGSVLAYYWSE FSIPQGULVEALERVMAEREVVMLPPARSILKS FSVPGULVEALERVMAEREVVMLPPARSILKS FSVPGULVEALERVMAEREVVMLPPARSILKS FSVPGULVEALERVMAEREVVMLPPARSILKS FSVPGULVEALERVMAEREVVMLPPARSILKS FSVPGULVEALERVMAEREVVMLPPARSILKS FSVPGULVEALERVMAEREVVMLPPARSILKS FSVPGULVEALERVMAEREVVMLPPARSILKS FSVVTSVVAPTDSKTVQRTODNSCSFGLHAR GVELMRFTTIGFDSFYPAHARCQWALRGD ADSVLSILTRSFDLASCDERGRHLVVTYYNTU. LITLITITITERFINGVEATFGLIPFRMSSCGGRL RAQGTFNSPYYPGHPYPNIDCTWHEVPNN QHYKVRFKFYLLEPGYPAGTCFKDVYENG EYCGERSQFVVTSNSKHTVRHSDSYDTOT GFLAEYLSYDSSDPCPGGFTCRTGRCRKELR CDGWADCTDHSBELNCSCDAGHOFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPAQTF RCSNGKCLSKSQCONGKDCGDOSDEASCP KNVVTVTCRHTYRCLNGLCLSKGPPECDGK EDCSDGSDBKDCDGCLRSFTRQAKVVGGTD ADGEGWPWQVSLHALGQGHLCGASLSPNWL VSAAHCVIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVGRRLKRISHFPRDHTTPDYDLALL ELEKPASTSSNVRPICLPDASHVFPAGKAIWV VSAAHCVIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVGRRLKRISHFPRDHTTPDYDLALL ELEKPASTSSNVRPICLDASHVFPAGKAIWV TOWGHTGYGGTGGALLQKGEIRVNDGTCEN LLPQGTTRMMCVGFLSGGVDSCQGDSGGFL LLPQGTTRMMCVGFLSGGVDSCQGDSGGFL LLPQGTTRMMCVGFLSGGVDSCQGDSGGFR LLPQGTTRMMCVGFLSGGVDSCQGDSGGFL LLPQGTTRMMCVGFLSGGVDSCQGDSGGFL LPQGTTRMGVKSQFCCAWS SSGTTKSGKATIRLYLKRNFDQEPLAKEVSL EQGIVLPRILEPRIJCHEDGFTQVTXNKFVLILVCKAV PATOFFFCNGEWVRQVBHVERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSGKATIRLYLRKNFDQEPLAKEVSL EQGIVLPCRPPEGIPPAL 450 1800 A 3620 1 2676 MPSLGGGMDLTCFFGVSPACGAQSWSIFG ADAAEVFGTRGISQQCAAMPHPDEDEPPGG PQAAQSAGQQGFTAGDFMAGNAGGFAGEP MGPEAGSKAELPFTVSRPPLLRGLSWDSQFE PGPRLQKVLAKLPLAEEKRFAGRAGGCLAK PYTNSSGKMALNSPQPGPVESELGKCN LALLEEAKLVSERTLTRGGKSSSPGDSFSA VSPLLSFSASFTSSSSSTSLTYPTPEGDEDAVS SPHGEFTNYKGLADKQCHTAFRAGSCSPTPTIVLGDA PPPTVSSGKMALNSPQFQFVESELGKCN LALLEEAKLVSERTLTRGGKSSSPGDSFSA VSPLLSFSASFTSSSSSSTSLTYPTPEGDEDAVS SPHGEFTNYKGLADKQCHTAFRAGSCSPPTIVLGDA PPPTVSSGKMALNSPQFQFVESELGKCN LALLEEAKLVSERTLTRGGKSSSSPGDSFSA VSPLLSFSASFT		İ					
PSSRSAPGLASGPARPLLQGYMSSAGPLLCYI			Ì	ļ	ļ		1 =
1798 A 3604 3115 557 FRRKGGGGPKPGAGLKYNSRIEKYNGLEE GVEFLPYNNYKKVEKHGPGRWYNLAAVLIGLEL LLI-VILGIGFLWWH, QYRDYRGV KYPRGYM KITNENFVDAYENSNSTEFVSI. ASKVEDJAL LYSGVPFLGPYHKESAYTAPSGSVIAYYWSE FSIPQHLVEEAERVMALERVWH, PRANSIKS FVYTSVVAFTIDSKTVQRTQMSCSFGLHAR GVELMRFTTIGFPDSYTPAHARCQWALRGD ADSVLSLTFRSFDLASCDERGRHLVTVYTYLL SPMEPHALVQLCGTYPPSYNLTHSISKONVL LITLITINTERHPGFEATFFQLPRMSSCGGL RKAQGTNSRYYPGHPPRINDCTWNIEVPSIN QHVKVRFKFFYLLEGVPAGTCRKDYVEIN GFLAFYLSYDSSDPCPGGFTCRTGRCIRKEL CIGWADCTHJSDELNSCDAGHGFTCKNKF CKPLEWVCDSLINLOGINSDEQGCSCPAQTF RCSNGKCLSKSQCNKKDDCGGSDEASCR CKVNVYTCTKHTYRCLNGLCLSKGNPECDGS EDCSDGSDEKDCCCQCAGTF CKNGKCLSKSQCNKKDDCGGSDEASCR CKVNVYTCTKHTYRCLNGLCSKSRYPECDG QRSAFGVQERAKRISHFFNDGYTDATAL LELEKPASYSSMVPEICLPDASHVPFAGKATUR VSAARICYIDDRGRTSSDPTGWTAFLCLHDOG GRAGGVQERAKRISHFFNDGXTICH TURQHTPRMCVGGLSGVDSCQGSCPAQTF CKNGKCLSKSQCNKKDDCGGSDEASCR CKVNVYTCTKHTYRCLNGLCSKSRYPECDG CROSSPOWENGCAGNLKGCASLISPWIL VSAARICYIDDRGRTSSDPTGWTAFLCLHDOG GRAGGVQERAKRGVYSMDGCAGNKRGVT TURQHTYNGCALLIQKGEINYGCT CKNGKCLSKSQCNKKDCCCC CKNGKCLSKSQCVSCDGCAGNKGCT CTRUFT CKNGKCLSKSQCVSCDGCAGNKGCT CTRUFT CKNGKCLSKSQCVSCDGCAGNKGCT CTRUFT CKNGKCLSKSQCVSCDGCAGNKGCT CTRUFT CKNGKCLSKSQCVSCDGCAGNKGCT CTRUFT CKNGKCLSKSQCVSCDGCAGNKGCT CTRUFT CKNGKCLSKSQCVSCDGCAGNKGCT CTRUFT CKNGKCLSKSQCVSCDGCAGNKGCT CTRUFT CKNGKCLSKSQCVSCDGCAGNKGCT CTRUFT CKNGKCLSKSQCVSCDGCAGNKGCT CTRUFT CKNGKCLSKSQCVSCDGCAGNKGCT CTRUFT CKNGKCLSKSQCVSCDGCAGNKGCCCCCC CKNGCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC		ĺ		1	1		
448 1798 A 3604 3115 557 FREKGGGEREDFAGALKYNSREEKYNGLEG GVEFLPVNINVEKVEKHOPGRWAVLAAVLIG LLI-VILLGGELVWHLQYRDVRVQKVFNGYM RITNENFVDAYENNSTEF-VSLASK-KUALUL LYSGYPFLGYPHESS-VTAFSES-SVLAY-YWSE FSIFQHL-VEEAER-WMAEER-VML.PPRABLSL FVYSVVAFFIDS-KTVQKTQDNSS-SGLHAR GVELMEPTTFGFPDSPYPAHAR-CQWALRGD ADS-VIS-TIRS-FSID-KSCDEGREHLVITV-YNTUL SPMEPHAU-VQL-CGTYPPS-NILTH-SISSONLY LITLI-ITMERHEPOER-AST-CEREBRI-VITV-YNTUL SPMEPHAU-VQL-CGTYPPS-NILTH-SISSONLY LITLI-ITMERHEPOER-AST-CEREBRI-VITV-YNTUL SPMEPHAU-VQL-CGTYPPS-NILTH-SISSONLY GFLAE-YLS-YDS-DCP-GQFT-CTR-GREEN GVELMEPTT-GER-TS-CL-PRASC-CGRL RKAQGTFNSPY-YGH-YPPNIDCT-WINE-VPNIN QH-KW-FREFF-LL-ER-OYA-GT-CH-LI-CH-DC-S-GWAD-CTH-ISDEL-IN-SCD-AGRIQFT-CKN-K- CKPLF-W-CDSL-IND-CGD-INS-DQG-CS-CP-AQTI- GFLAE-YLS-YDS-DCP-GQFT-CKN-K- CKPLF-W-CDSL-IND-CGD-INS-DQG-CS-CP-AQTI- GFLAE-YLS-YDS-DCP-GQFT-CKN-K- CKPLF-W-CDSL-IND-CGD-INS-DQG-CS-CP-AQTI- GFLAE-YLS-YDS-DCP-GQFT-CKN-K- CKPLF-W-CDSL-IND-CGD-INS-DQG-CS-CP-AQTI- GRAP-GV-GG-CR-ST-GRA-W-VGGTD ADEGEWPWQV-SLHALG-GGHL-CAS-LIS-PR-IN- ADEGEWPWQV-SLHALG-GGHL-CAS-LIS-PR-IN- LI-CAS-CO-CG-ST-ST-GAR-W-VGGTD ADEGEWPWQV-SLHALG-GGHL-CAS-LIS-PR-IN- LI-CAS-CO-CG-ST-GR-GR-GR-GR-GR-GR-GR-GR-GR-GR-GR-GR-GR-	1			ĺ		1	
GVEFLPVNNYKKVEKHOPGRWYVLAAVILIG LLI-VLLGIGFL-WHLQYRDVRVQKVYNGYM RITSENFVDAYENSNSTEFVSLASK-VKDALKL LYSGVPFLGFYHKESAVTAFSEGSVIAYYWS SIPIQHI-VEEARFWAEREV-WAMPPRARSILKS FVVTSVVAFFIDSKTVQRTQDNSCSFGLHAA GVELMRFTTFOFDFSYPHAFAGCQWALRGD ADSVLSLTFRSFDLASCDERGRHL-VIT-VTHL SPMEPHAU-VQLCGTYPPSYNLTFISSIONVL LITILITINTERRHPGIFEATFFQLPRMSSCGGT, RKAQGTFNSPY-VGHYPPHIQTCTWHEVPNN QHVKVRFKFFYLLEPGVPAGTCPKNYVEING EKYCGERSQFV-VTSNSNKITVRFHSDSSYTDT GFLASTJSTNSSDPCPGGFTCKTGRCIRKEILR CDGWADCTDHSDELNCSCDAGHGFTCKNFF KCSNCKCLSKSQQCNCKDDCGDGSDEASCP KVNVVVTCKHTYRCHNGLLSKGPCCSCP-AQTT RCSNCKCLSKSQQCNCKDDCGDGSDEASCP KVNVVVTCKHTYRCHOLGLSKGPCCGCSCP-AQTT RCSNCKCLSKSQQCNCKDDCGDGSDEASCP KVNVVTCKHTYRCHOLGLSKGPCCGCSCP-AQTT RCSNCKCLSKSQQCNCKDDCGDGSDEASCP KVNVVTCKHTYRCHOLGLSKGPCCGCSCP-AQTT RCSNCKCLSKSQQCNCKDDCGDGSDEASCP KVNVVTCKHTYRCHOLGLSKGPCCGCSCP-AQTT RCSNCKCLSKSQQCNCKDDCGGSDEASCP KVNVVTCKHTYRCHOLGLSKGPCCGCSCP-AQTT RCSNCKCLSKSQQCNCKDDCGGSDEASCP KVNVVTCKHTYRCHOLGLSLGSCDSCCSCP-AQTT RCSNCKCLSKSQQCNCKDDCGGSDEASCP KVNVVTCKHTYRCHOLGLSCLSCDSCCSCP-AQTT RCSNCKCLSKSQQCNCKDDCGGSDEASCP KVNVVTCKHTYRCLNCLSKGPCCGSCGSCGPL SVSAAHCYDDRGFRYSDPTQWTAFLCHDQS QRSAGVQERRLKRISHFFRNDFTDTDVTDLL LELEKRAFSSSMVEPICLPDASHTYPAGKAHW TOWGHTYVGGTGALLLQKGEIRVNQTTCCL LLPQQTTPRMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKFGYY TLLFLFDWIKENTGV FXLFFDWIKENTGV FXSGFPRSTMOGTSTERLEARRPAGRIPWSSGC ANPDLLPHFLGFPEDVYVXCNCPVLLVCKAV PATQIFFKCNGGWWSAQVGSATVANPVPG ANPDLLPHFLGFPEDVYVXCNCPVLLVCKAV PATQIFFKCNGGWWSAQVGSATVANPVPG ANPDLLPHFLGFPEDVYVXCNCPVLLVCKAV PATQIFFKCNGGWWSAQVGSATVANPVPG ANPDLLPHFLGFPEDVYVXCNCPVLLVCKAV PATQIFFKCNGGWWSAQVGSCPPTTVLTCAG TYPEGGETDKILAMRVSSPHCRLSRHLKVSI ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSTORGSTGAGGPAGFP MGPEGTSNKAMRVSSPHCRLSRHLLKVSI ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSTORGGGPAGFP MGPEGTSNKAMRVSSPHCRLSRHLLKVSI ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSTORGGGPAGFP MGPEGTSNKAMRVSSPHCRLSRHLLKVSI ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLIKASHTSSAGRAGGKAA PGPLXDPOJOVPRVGLLARDRYNGGRPAGFP MGPEGASSAALIPTVSRPLLEGGSLSP	140	1500	 	0.604			
LLI-VILGIGFL/WHILQYRDVR-VOKYN-TOWN RITNEWN-PUAY-RINSNETF-YSLS-KVENDALKIL LYSGVPFLGPYHKESA-VTAFSEGSVIA-YYWSE FSIPQHLVEE-AREWAMEER-VVMLPPRARISL FVVTSVV-AFFIDSKIT-VQRTQDINSCSFGLHAR GVELMRFTTGFPDSYP-AHARCQWALRGD ADSVLS-LITRESTDLASCDEGGRIH-NITV-YYNTL SPMEPHAU VQL.CGTYPPSYNLTH-SISSONU-LITLLITH-THREHPOFPAGT-PFAHSCGGRIL RK-AQGTP-NSPY-YFGH-PPPRIDCT-WINE-PPNN QH-WYR-PFK-PTL-HEPOV-PAGT-PC-WENSCGGRIL RK-AQGTP-NSPY-YFGH-PPRIDCT-WINE-PPNN QH-WYR-PFK-PTL-HEPOV-PAGT-PC-WENSCGGRIL RK-AQGT-NSPY-YFGH-PPRIDCT-WINE-PPNN QH-WYR-PFK-PTL-HEPOV-PAGT-PC-KCK-K-K-CKPL-W-WCDSL-NDC-GDNSDEQ-GCS-CP-AQTF RCSN-MCK-LSSK-QCN-GKD-DCG-GSDE-ASC-C-KV-NV-VTC-WHI-YR-CL-NOLCL-SK-GN-PE-CDG-EK-YC-W-W-W-SK-LSK-G-CK-LSW-PC-QC-G-K-ST-W-R-W-G-G-G-CS-CP-AQTF RCSN-MCK-LSSK-QCN-GKD-DCG-GSDE-ASC-C-KV-NV-VTC-WHI-YR-CL-NOLCL-SK-GN-PE-CDG-G-CK-ST-W-R-W-G-G-G-CK-W-W-W-W-W-W-W-W-W-W-W-W-W-W-W-W-W-W-	448	1798	Α	3604	3115	557	
### A ### A			l				
LYSGVPFLGPYHIKESAYMAPRASEGSVALYWSE FSIPOHLVERARWAMERRAYMAPRASIKS FFVTSVVAFPTDSKTVQRTQDNSCSFGLHAR GVELMRPTTPGFPDSPYPAHARCQWALRGD ADSVLSLIFRSFOLASCDERGRHLVTVYNTLL SPMEPHALVQLGTYPFSYNLTHSSSQNVL LITILITNTERRHGPGFATFEQFLPRMSSGGGRL RKAQGTFNSPYYPGHYPPNIDCTWNIEVPNN QHYKYRKFFYLLEPGVPAGTCKDYVENG EKYGGERSOFVYTSNSKITVRFHSDQSYTDT GFLAEYLSYDSSDPCPGQFTCRTGRCRERCR CDGWADCTDHSDELNSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEGGCSCPAQTF RCSNGKCLSKSQCNGKDDCOSDEASCP KVNVVTCTKHTYRCLNGLCLSKGPPECDGK EDCSGSBEKDCDCGLSFFFANDTGTCKNTV ADEGEWPWQVSLHALGQGHIGGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPQVQERRLKRISSFFFNDFTFDVDIALL ELEKPAEYSSMVRPICLPDASHVPPAGKANW TGWCHTQYGGTGALLJQKGERLVINQTTCEN LLPQQTTPRMMCVGFLSGGVDSCQGDSGGPL SSVADGRIPQAGVVSWGDGCQRNKRGVY TRIPLFRDWIKENTGV FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRFSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEFEDVYIVKNKPYLLVCKAV PATQIFFKKNGEWVRQVDHVIRRSTDGSSGLP TMEVRINVSRQQVEKVFGLEFYWCQCVAWS SSGTTKSGKAYKRIAYLRKNFSQEPLAKEVSL EQGIVLPCRPFGGIPPAE 450 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPTAGVSSCGDFSKKN LALLEEAKLVSERTLTRRGRKRSSSFGDSPSA VSPNLSPSASPTSSRSNLTVPTPPEGDEADVS SPPNLSPSASPTSSRSNLTVPTPPEGDEADVS SPNLSPSASPTSSRSNLTVPTPPEGDEADVS SPPNLSPSASPTSSRSNLTVPTPPEGDEADVS ARCHITTAGAT AUGH				ļ			1
FSIPQHLVEEARRYMAERRYVMLPPRARSILKS FVVTSVVAPPTDSKTYORTODSCF0GHAR GVELMRPTTPGFPDSPYPAHARCQWALRGD ADSVLSLTFRSFDLASCDERGRHLVITVYNTUL SPMEPHALV QUCGTYPPSYNLTHSISQINVL LITLITINTERRHPGFEATFFQLPRMSSCGGRL KAQGTINSPYPGHYPPNIGHYPMISSQCGRL RKAQGTINSPYPGHYPPNIGHYPMISSQCGRL RKAQGTINSPYPGHYPPNIGHYPMISSQCGRL RKAQGTINSPYPGHYPPNIGHYPMISSQCGRL RKAQGTINSPYPGHYPPNIGHYPMISSQCGRL GKYVENTSNSKHITVRFTSQCTSCKERER CDGWADCTDHISDELNCSCDAGHGTICKKF CKPLFWCDSLDMCGGMSDEGGSCSPAQTIT RCSNGKCLSKSQCORKDDCGDGSDEASCP KVNVTCTKHTYRCLNGLCJKGMPECDGK EDCSDGSDEKDCDCGGRSDEGGSCSPAQTIT RCSNGKCLSKSQCORKDDCGGSGSGCPAQTIT RCSNGKCLSKSQCORKDDCGGNSGDGCGSCSTAQTIT RCSNGKCLSKSQCORKDDCGGNSGGGSCCPAQTIT RCSNGKCLSKSQCORKDDCGGNSGGGSCSCPAQTIT RCSNGKCLSKSQCORKDDCGGNSGGTC ADEGUPWQVSLHALOQGHICGASLISPNUL VSAAHCYIDDRGFRYSDPTQUTAFLGLHDQS QRSAPCVQERLKRIISPFFNDF1FDVIDALL ELEKPAEYSSMXPRICLPDASAVPPQAGAANV TGWGHTOYGGTGALLLQKGEIRVINQTTCEN LLPQQTTPRMCVQFLSQGDSGGPL SSVEADGRIPQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLPHFLGGFEDDYVINVPLU/CKAV PATQFFFKCNGGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQVEKVFGLEEYWCCVAWS SSGTITKSQKAYLRIAYLRKNFEGEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MPPSLOGGMDDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQCEAAMPHIPEDEEPPGE PAQASPAGQOPPTAGVSCSTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEBCKKN LALLEEAKLVSERLTTRGRISKSSSGDSPSA VSPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SPNLSPSASPTSSRSNLTVPTPPEGDADVS SSPNLSPSASPTSSRSNLTVPTPPEGDADVS ALLEEALERANCHEMT ALLEEALLVSA							
FVVTSVVAPTDSKTVQRTQDNSCSFGLHAR GVELMRTITGPFDSYPYAHCQWALRGD ADSVLSLTFRSFDLASCDERGRHLVITVYNTUL SPMEPHAUVQLGGTTYPSYNLTPHSSIGNVL LITILITMTERRHPGVEATFFQLPRMSSCGGRL RKAQGTFNSPYYPGHYPPNDCTWNIEVPNN QHYKVRFKFYLLEPGVPAGTCFKDYVENN EKYCGERSOFVVTSNSKITVRFHSDQSYTDT GFLAEYLSYDSSDPCPGQFTCRTGRCRERCR CDGWADCTDHSDELNSCLOAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPAQTF RCSNGKCLSKSQCNGKDDCGSDEASCP RCSNGKCLSKSQCNGKDDCGSDEASCP RCSNGKCLSKSQCNGKDDCGSDEASCP RCSNGKCLSKSQCNGKDDCGSDEASCP RCSNGKCLSKSQCNGKDDCGSDEASCP RCSNGKCLSKSQCNGKDDCGSDEASCP RCSNGKCLSKSQCNGKDDCGSDEASCP RCSNGKCLSKSQCNGKDDCGSDEASCP RCSNGKCLSKSQCNGKDDCGSDEASCP RCSNGKCLSKSQCNGKDDCGSDEASCP RCSNGKCLSKSQCNGKDDCGSDEASCP RCSNGKCLSKSQCNGKDDCGSDEASCP RCSNGKCLSKSQCNGKDCGSDEASCP RCSNGKCLSKSQCNGKDCGSDEASCP RCSNGKCLSKSQCNGKDCGSDEASCP RCSNGKCLSKSQCNGKDCGSDEASCP RCSNGKCLSKSQCNGKDCGSDEASCP RCSNGKCLSKSQCOKGDCSGCAATANFORD VSAAHCTUDDRGFRYSDPTQWTAFT.ELHOG RGSHTYNOTTCHUD LEEKPAEYSSMYRPICLPDASHVPFAGKAHW TGWGHTQYGGTGALILQKGERVNQTYTCHED LEEKPAEYSSMYRPICLPDASHVPFAGKAHW TGWGHTQYGGTGALILQKGERVNQTYTCHED LEEKPAEYSSMYRPICLPDASHVPFAGKAHW TGWGHTQYGGTGALILQKGERVNQTYTCHED LEEKPAEYSSMYRPICLPDASHVPFAGKAHW TGWGHTQYGGTGALILQKGERVNQTYTCHED LEEKPAEYSSMYRPICLPDASHVPFAGKAHW TGWGHTQYGGTGALILQKGERVNQTYTCHED LEEKPAEYSSMYRPICLPDASHVPFAGKAHW TGWGHTQYGGTGALILQKGERVNQTYTCHED LEEKPAEYSSMYRPICLPDASHVPFAGKAHW TGWGHTQYGGTGALILQKGERVNQTYTCHED LEEKPAEYSSMYRPICLPDASHVPFAGKAHW TGWGHTQAGVAWAGAATANFYPGA ANPDLLPHFLGEPDVYIVNRFVLLVCKAV PATQFFRCNGGWVRQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYRHAYLRKNFEQEPLAKEVSL EQQILPCRTPPGIPAE ADAAEVPGTRGHSQQEAAMPHIFEDEEPPOE PQAAQSPAGQQFPTAGVSCSTPTTVLTGDA TSPEGETUKNLANRVFSPHKRLSIRHLKVST ASLTSVPBAGHTBULVNDQLPGADVS SSPHLSPSASPTSSRSNSLTVPTPFEGDABAVS SSPHLSPSASPTSSRSNSLTVPTPFEGDABAVS SSPHLSPSASPTSSRSNSLTVPTPFEGDABAVS SSPHLSPSASPTSSRSNSLTVPTPFEGDABAVS SSPHLSPSASPTSSRSNSLTVPTPFEGDABAVS SSPHLSPSASPTSSRSNSLTVPTPFEGDABAVS SSPHLSPSASPTSSRSNSLTVPTPFEGDABAVS SSPHLSPSASPTSSRSNSLTVPTPFEGDABAVS SSPHLSPSASPTSSRSNSLTVPTPFEGDABAVS SSPHLSPSASPTSSRSNSLTVPTPFEGDABAVS SSPHLSPSASPTSSRSNSLTVPTPFEGDABAVS SSPHLSPSASPTSSRSNSLTVPTPFEGDABAVS SSP	•			ĺ		{	
GVELMERTTPGFPDSYPYAHARCQWALRGD ADSVLSLIFESFDLASCDERGRHLWTVYNTIL SPMEPHAL VOLGGTYPPSYNLTFHSISONVL LITLITNTERRHPGFEATFFQLPRIMSSCGRI RKAQGTFNSPYYPGHYPPNIDCTWNIEVPNN QHYKVRRKFFYLLEFGYPAGTCPKDYVEING EKYCGERSGPVYTSNSKRITYBDQSYTDT GFLAEYLSYDSSDPCPGQFTCRTGRCRELR CDGWADCTDHSDELNCSDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPAGTT RCSNGKCLSKSQQCNGKDDCDGSDEASCP KVNVVTCTKHTYRCLNGLCJKGNPECDGK EDCSDGSBEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGICASLISPNWL VSAAHCYDDRGFRYSDPTQWTAFLCLHDQS QRSAPGVQERLKRIISHPFFNDFTDOTLALL ELEKPAEYSSMYRPICLPDASHVPPAGKAIWV TGWCHTQYGGTGALLLQKGERVINGTTCEN LLPQGTPRMKCVGFLSGGVDSCQGDSGGPL SSVEADGRIPQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRSLMAGRQHGWSAQQSATVANPVPG ANPDLPHFLGEFPDYVINVFLVLCKAV PATQFFKCNGGWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEFFWCYCVAWS SGGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQQIVLPCRPPEGIPPAG ADAAEVPGTRGHSQQEAAMPHIFEDEEPPGE PAAQSPAGQOQFPTAGVSCSTPTIVLTGDA TSPEGETUKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERTLTRGRKSRSSFGODSSP SPHLPSTSSRSNSLTVPTPFEGDADVS SPHLPSSASPTSSRSNSLTVPTPFEGDADVS SPHLPSSASPTSSRSNSLTVPTPFFGLAA PVTNSSGKMALNSPQGPPVESLEKQLLKTG WEGSPLRSSTTQDAAGVGPPASQGRAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPE PGPRLQKVLAKLPLAEEEKFRAGKAGGKLAK APCLKDFQIQVQPVRMQKLTKLREHHLMRN QNLVGLKLPDLSEAAEGEGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHHLMRN QNLVGLKLPDLSEAAEGEGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHHLMRN QNLVGLKLPDLSEAAEGEGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHHLMRN QNLVGLKLPDLSEAAEGEGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHHLMRN QNLVGLKLPDLSEAAEGEGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHHLMRN QNLVGLKLPDLSEAAEGEGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHHLMRN QNLVGLKLPDLSEAAEGEGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHHLMRN QNLVGLKLPDLSEAAEGEGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHHLMRN QNLVGLKLPDLSEAAEGEGKAGGKLAK APGLKDFQIQVQPVRAGGRAGGFLAG RERNLTEENTEKSLENKASISSASSLWHHCE BERNLTEENTEKSLENKASISSASSLWHHCE BERNLTEENTEKSLENKASISMSASLWHHCE				1			FSIPQHLVEEAERVMAEERVVMLPPRARSLKS
ADSVLSLTFRSFDLASCDERGRHL'ATVTYNTUL SPMEPHALVQLCGTYPESYLTFHSISQNVL LITLITITERHPGIFEATFFQLPRMSSCGGL RKAQGTFNSPYYPGHYPRIDICTWISEVPRN QHYKVPRKFYTLLEGVPAGTCPKDYVENIG EKYCGERSQFVYTSNSNKITVRFHSDGSYTDT GFLAEYLSYDSSDPCPGCFTGRCIRKELR CDGWADCTDHSDELNCSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPAQTIF RCSNGKCLSKSQQVCNKGDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSBEACDCCGLTSKRQARVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGGRYSDPTQWTAFLGLHDQS QRSAPGVQERLKRIISHPFNDFTFDYDIALL ELEKPAFYSSMVPPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALLLQKGERIVNQTTCEN LLPQQITPRMMCVGFLSGQVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGYY TRLPLFEDWIKENTGV 449 1799 A 3618 2 613 PVSGSPWRMDGSTERLEARRPAGRIPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEFWCQCVAW SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPEGIPPAG ADAAEVFGTKGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQGGPTAGVSCSPTETIVLTGDA TSPEGETDKNLANRVHSPHKRLSTRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERPLTRGRKSRSSPGDSPSA VSPNLSPSASSTSSRSNSLTYFTPEDDEADVS SPHGEFDVYNKGLADRKQNDQKKVSQGRLAP RPPPVEKSKEIAHEQKENFDPLQYPETTPKGLA PVYNSSGKMALNSPPQPESELGKQLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPFTVSRFPLLRGLSUNDSCPEE PGFRLQKVLAKLPLAEEEEKERAGKAGGKLAK APGLKDFQIQVQPVXMQKLTKLREEHLIMRN QNLYGLKLPLAEEEKERAGKAGGKLAK APGLKDFQIQVQPVXMQKLTKLREEHLIMRN QNLYGLKLPLAEEBLENKRK QNLYGLKLPLAEEBLERLEMENK QNLYGLKLPLAEEBLERLEMENK QNLYGLKLPLAEEBERFGGERPSLSPAHEEE ESKSGLDVMPNISDVLLKKRVHRSLPGSAPP LTEKEVENVPVQLSSAFRNDSYTLESRINQAE RERNLTEEMTEKELENTKKASTISSASLWHHCE	, ·			l		, ·	FVVTSVVAFPTDSKTVQRTQDNSCSFGLHAR
SPMEPHAL VQL.CGTYPPSYNLTFHSSQONTL LITLITNTERRHPGIFEATFFQLPRMSSCGGRL RKAQGTFNSPYYPGHYPPNIDCTWMIEVPNN QHYKVRFKFYLLEPGVPAGTCPKDYESING EKYGGERSQFVYSTSNSHTVRFHSDQSYTDT GFLAEYLSYDSSDPCPGQFTCRTGRCIRKELR CDGWADCTDHSDELNCSCDAGHGPTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPAQTT RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCKHTYRCLNGLLSKGNPECDGK EDCSDGSDEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHLGGASLISPNWL VSAAHCYTDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERLKRIISHPFNDFTPDVIALL ELEKPAEYSSMVEPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITRMMCVGFLSGGVDSCQDSGGPL SSVEADGRTGQAVVSWGDGCAQRNKPGYY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMIGGSTERLEARRPAGRLPWSSRQ EMTRRFSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVNNFVLLVCKAV PATQIFFKCNGEWVRQVDHYERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPAKEVSL FQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGGGMDLTCFFGVSPACGAQASWSFG ADAAEVFGTRGHSQQCAAMPHIPEDEEPPGE PQAAQSPAGQQGPTAGVSCSPTPTIVLTGDA TSPEGETIDKINLANRYHSPHKRLSTRHLKVST ASLTSVDPAGHIDLVNDQLPDISISSEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSFGDSPSA VSFNLSPSASPTSSRSNSLTVPTPFPEGDEADVS SPHPGEFNVPKGLADRKQNDQRKVSQGRLAP PVTNSSGKMALNSPQPGPYSEGLGKQLLKTG WEGSPLPRSFTQDAAGVGPPASQGRFAGEP MGPFALGKVLAKLPLAEEEKFRAGKAGGKLAK APGLKDFQIQVQPVXMQKLTKLREREILMRN QNLYGIKLPDLSEAAEGEKGLPSTLSSPAIEE ESKSGLDVMPNISDVLLRKLRVIRRSLPGSAPP LTEKEVENVPVQLSSAFRNDSYTLESRINQAE EEKNLTEENTEKELENIKKSSTISSASLWHHCE							
LITLITNTERRHPGIFEATFFQLPRMSSGGRL RKAQGTPNSPYYPGHYPPNIDCTWNIEVPNN QHYKVRKFFYLLERGVAGTCPKDYYENG EKYGGERSQFVYTSNSNKITVRFHSDQSYTDT GFLAEYLSYDSSDCPGGFGCSCPAQTF RCSNGKCLSKSQQCNGKDDCGDCSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSBERDCDCGIRSFTRQARVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNUL VSAAHCYDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERLKRIISHPFFNDFTFDVIDLIL ELEKPAEYSSMYRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALLQKGEIRVINQTTCEN LLPQQTTPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSFWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIENTGSCVAW SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGVLVPCPPEGIPPAG 450 1800 A 3620 1 2676 MEPSLGGGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPTAGYCSSPTFIUTJGDA TSPEGETDKNLANRVHSPHKRISTSHLKVST ASLTSVDPAGHIDLVNDQLPDISISEBCKKN LALLEEAKLVSER-LTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPEDDEADVS SPHGEFNVPKGLADRKQNDQRKVSQGRLAP RPPPVKSKEIAIEQKENFDPLQYPETTPKGLA PVYNNSGKMALNSPQPGPVSELGKQULKTG WEGSPLPRSFTODAAGVGPPASQGRGPGP MGPEAGSKAELPPTVSRPPLLGCLSWDSGPEE PGFRLQKVLAKLPLAEBEKRRAGKAGKCHAK APGLKDFQIQVQPVXMQKLTKLREREILMRN QNLYGISLIPDLSEAAEDGEGLPSSLSPABEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVPVQLSSAFRNDSYTLESRINQAE RERNLTEEMTEKELENIKASITSSASLWHLCE						ł	ADSVLSLTFRSFDLASCDERGRHLV\TVYNT\L
RKAQGIFNSPYYPGHYPRIDCTWNIEVPNN QHYKVRFKFYLLEPGVPAGTCPKDYVEING EKYCGERSQFVTSNSNKITVRFHSDQSYTDT GFLAEYLSYDSSDPCPQGFTCRTGRCERKELR CDGWADCTDHISBELNGCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEQGGSCPAQTF RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLIGLLSKGNPECDGK EDCSDGSBEKDCDCGLRSFTRQARVGGTD ADEGEPWQVSLHALGQGHCGASLISFNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAFGVQERRLKRIISHFFNDFTFDYDLALL ELEKPAPSYSMVPRICLPDASHVFPAGKAIWV TGWGHTQYGGTGALLLQKGEIRVINQTTCEN LLPQQITPRMCVGFLSGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSSGPWBMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHTLGPEDVYIVKNKVLLVCKAV PATQIFFKCNGEWVRQVBHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEPWCQCVAWS SSGTTKSQKAYRIAJVRKNFFQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQCAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPITTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSSSPGDSPSA VSPNLSFSASPTSSRSTLTYPTPFGDBADVS SPHEGEPNVFKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEJAIEQKENFEDLGYPETTFKGLA PVTNSSGKMALNSPQFPVESELGKQLLKTG WEGSILPRSTTQDAAGVCPPASQGRGPAGEP MGPEAGSKAELPPTVSGLLATG WEGSILPRSTTODAAGVCPPASQGRGPAGEP MGPEAGSKAELPPTVSGLLATG WEGSILPRSTTODAAGVCPPASQGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGRGPAGEP MGPEAGSKAELPPTVSAGGRGPAGEP MGPEAGSKAELPPTVSAGGRGRGPTSLSPAMEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVPQLSSAFRNDSYTLESRINQAE RERNLTEENELNKASITSSASLWHHCE	1		ĺ			<u> </u>	SPMEPHA/LVQLCGTYPPSYNLTFHS/S/QNVL
OHWKYRFKFYJLERGVPAGTCPKDYVEING EKYCGERSQFVVTSNSNKITVRFHSDQSYTDT GFLAEYLSYDSSDPCPGQFTCRTGRCIRKELR CDGWADCTDHSDELNCSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEGGSCPAQTF RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLINGLCLSKGNPECDGK EDCSDSGBEKDCDCGLGRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERLLKRISHFFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHFPAGKAIWV TGWGHTDYGGTGALLQKGEHRVINQTTCEN LLPQQITPRMMCVGFLSGGVDSCQDDSGGPL SSVEADGRIPQAGVVSWDDGCAQRNKPGVY TRIPLFRDWIKENTIG SSVEADGRIPQAGVVSWDDGCAQRNKPGVY TRIPLFRDWIKENTIG A 3618 2 613						l	
EKYCGERSOFVYTSNSNKITVRFHSDQSYTDT GFLAEYLSYDSSDPCRGGTCRTGRCIRKELR CDGWADCTDHSDELNCSCDAGHGTCKNKF CKPLFWVCDSLNDCGDNSDEGGSCPAQTT RCSNGKCLSKSQQCGKCDDCGDSDBSASCP KVNVVTCTKHTYRCLNGLCLSKONPECDGK EDCSDGSDERDCDGCLRSFTRQARVGGTD ADEGEWPWQVSLHALGGGHLCGASLISFNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVOERLLKRISHFFRDFTFDYDTALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALLLQKGEHRYNQTTCEN LLPQQITPRMMCVOGFLAGVDSCQGBGGPL SSVEADGREPQAGVVSWGDGCAQRNKPGVY TRLPLFRMMCVGFLSGGVDSCQGBGGPL SSVEADGREPQAGVVSWGDGCAQRNKPGVY TRLPLFRMMCVGFLSGGVDSCQGBGGPL SSVEADGREPQAGVVSWGDGCAQRNKPGVY TRLPLFRGMVRCHVINGNVPLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQVEKVFGLEERYWCQCVAWS SGGTTKSQVEKVFGLEERYWCQCVAWS SGGTTKSQVAYRIKJVRKNFFQLFLAKEVSL EQGIVLPCRPPEGIPPAE 450 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHPEDEEPPGE PQAAQSPAGQQGPTAGVSCSPTTIVLTGDA TSPEGETDKALANRVHSPHRLSFRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKN LALLEEAKLVSERTLTRRGRKSRSSPGDSPSA VSPHLSFSAPTTSSRSLTLYTPPFEGDEADVS SPHEGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKELAIEQKENFTPJSLCGL VSPHLSFSAPTSSRSLTLYTPTPEGDEADVS SPHEGERNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKELAIEQKENFTPJNSGFEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDPQIQVQPVRMQKLTKLREHLIMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELEHNKASITSSASLWHHCE			ĺ			İ	RKAQGTFNSPYYPGHYPPNIDCTWNIEVPNN
GFLAEVLSYDSSDPCPGOFTCKTGRCIRKELR CDGWADCTDHSDELNCSCDAGHQFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPAQTF RCSNGKCLSKSQQCNGKDDCGGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSDEKDCDCGLSSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLSPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERRLKRIISHPFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALLQKGEIRVNQTTCEN LLPQQTTPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRSILMAGRQWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVERSTDGSSGLP TMEVERNVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCFFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGGRKSRSSPGDSPSA VSPNLSPSASPTSSSNSLTVPTIPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKELIAEQRENDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSSPTQDAAGVGPPASQGGGPAAGP MGPEAGSKAELPTVSRPPLLRGLSWDSGPEE POPRLQKVLAKLPLAEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAAEQEKGLPSSLPAIEEE ESKSGLDVMPNISDVLLRKLRVIRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELEINFKASITSSASLWHHCE	}	1	ļ		•		QHVKVRFKFFYLLEPGVPAGTCPKDYVEING
CDGWADCTDHSDELNCSCDAGHOFTCKNKF CKPLFWVCDSLNDCGDNSDEQGCSCPAQTF RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNFECDGK EDCSDGSDEKDCDGRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERRLKRISHPFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALLQKGERVNDQTTCEN LLPQQITPRMMCVGLSGGVDSSCQDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV TRLPLFRDWIKENTGV ANPDLLPHFLGEFEDVYIVNEVLLVCKAV PATQIFFKCNGEWVQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPFEGIPTGFT ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASFTSSRSNSLTVPTPPEGDEADVS SPHGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKELAEQKENFDPLQYPETTPKGLA PYTNSSGKMALNSPQFQPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE				l .			EKYCGERSQFVVTSNSNKITVRFHSDQSYTDT
CKPLFWVCDSLNDCGDNSDEQGCSCPAQTT RCSNGKCLSKSQCNGKDDCGDGSDEASCP KVNVTCTKHTYTCLNGLCLSKGNPECDGK EDCSDGSDEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERRLKRIISHPFNDFTFDYDIALL ELKPAGYSSMVRPICLPDASHVFPAGKAIWV TOWGHTQYGGTGALLQKGEIRVNQTTCEN LLPQQTTPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIPQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEFEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVPGLEYWCQCVAWS SSGTTKSQKAYRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWISIFG ADAGEVPGTRGHSQQEAAMPHIPDBEEPPGE PQAAQSPAGQQGPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASFTSSRSNSLTVYPIPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP PVTNSSGKMALNSPQGFBLAP PVTNSSGKMALNSPQGPBVSGGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPE PGPRLQKVLAKLPLAEEKRFAGKAGGKLAK APGLKDFQIQVQVPVRQKLTKLREEHLMRN QNLVGLKLPDLSEAAEQEGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	i						GFLAEYLSYDSSDPCPGQFTCRTGRCIRKELR
RCSNGKCLSKSQQCNGKDDCGDGSDEASCP KVNVVTCTKHTYRCLNGLCLSKGNPECDGK EDCSDGSDEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHIGGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERELKRIISHPFFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHVPPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGVDSCQGDSGGFL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGGGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSIRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERPLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVKSKELAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENRKASITSSASLWHIGE	ĺ			[1	CDGWADCTDHSDELNCSCDAGHQFTCKNKF
KVNVVTCTKHTYŘCLNOLCLSKGNPECDGK EDCSDGSDEKDCDCGLRSFTRQARVVGGTD ADEGEWPWQVSLHALGQGHICGASLISPNWL VSAAHCYIDORGFRYSDPTQWTAFLGLHDQS QRSAPGVQERRLKRIISHFFNDFITDYDLALL ELEKPAEYSSWAVPICLPDASHVPPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARPAGGRIPWSSRQ EMTRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGITKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHIDLVNDQLPDISISEEDKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSLTVPTPPEGDEADVS SSHPGEFNVPKGLADRKQNDQRKVSQGRLAP RPPVVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSFTDDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDPQIQVQPVRMQKLTKLREHILMRN QNLVGKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRNQAE	1	ĺ	ĺ	į i		ĺ	CKPLFWVCDSLNDCGDNSDEQGCSCP\AQTF
EDCSDGSDEKDCDCGRSFTRQARVVGGTD ADEGEWPWQVSLHALGQHILGASLISPNWL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERRLKRIISHPFFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHVPPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVRNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPFEGETDIXNLANRVHSPHKRLSTRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEFDVFKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENIDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSTYDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDPQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENNYVQLSSAFRNDSYTLESRNQAE RERNLTEENTEKELENFKASITSSASLWHICE	'						RCSNGKCLSKSQQCNGKDDCGDGSDEASCP
ADEGEWPWQVSLHALGQGHICGASLISPNWIL VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS QRSAPGVQERRLKRIISHFFNDFTFDYDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALIQKGEIRVINQTTCEN LLPQQITPRMCVGFLSGGYDSCQGDSGGFL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHIGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSKRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHRGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQEENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSSPLRRSTYDDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVPYMMGKLTKLREHILMNN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRNOAE RERNLTEENTEKELENFKKSITSSASLWHHICE			ļ				KVNVVTCTKHTYRCLNGLCLSKGNPECDGK
VSAAHCYIDORGFRYSDFTQWTAFLGLHDQS QRSAPGVQERRLKRIISHPFNDFIFDYDIALL ELEKPAFYSSMYRPICLPDASHVPPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITTPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLPHFLGEPEDVYIVKNKPVLL VCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPFGGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQFPTAGVSCSPTFTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASFTSSRSINSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKELAIEQKENFDPLQVPETTPKGLA PVYINSSGKMALNSPQPGPVESELGKQLIKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESSRNQAE RERNLTEENTEKELENFKASITSSASLWHHICE	,			1			EDCSDGSDEKDCDCGLRSFTRQARVVGGTD
QRSAPGVQERRLKRISHPFFNDFTFDVDIALL ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMCVOFLSGOVDSCQGDSGGFL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTIGV 449 1799 A 3618 2 613 FVSGSPWRDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNRPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPTTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGKSRSSPGDSFSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVYINSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISIDVULRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE							ADEGEWPWQVSLHALGQGHICGASLISPNWL
ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVTERSTDGSSGLP TMEVRINVSRQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPTTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKN LALLEEAKLVSERFLTRRGKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHFGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKELAIEQKENFDPLQYPETTPKGLA PVYTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEREVENVFVQLSSAFRANDSYTLEESRNQAE RERNLTEENTEKELENFKASITSSASLWHHCE		f i					VSAAHCYIDDRGFRYSDPTQWTAFLGLHDQS
TGWGHTQYGGTGALILQKGEIRVINQTTCEN LLPQQITPRMMCVGFLSGGVDSCQGDSGGFL SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGGGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSFTPTTVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPTVYKSKEJAFQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKACHELPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVEVNVFVQLSSAFRNDSYTLESRNQAE RERNLTEENTEKELENFKASITSSASLWHHCE		i I					QRSAPGVQERRLKRIISHPFFNDFTFDYDIALL
LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGYY TRLP1.FRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFK.CNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAG 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKONDQRKVSQGRLAP RPPPVEKSKELAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRNQAE RERNLTEENTEKELENFKASITSSASLWHHCE							ELEKPAEYSSMVRPICLPDASHVFPAGKAIWV
LLPQQITPRMMCVGFLSGGVDSCQGDSGGPL SSVEADGRIFQAGVVSWGDGCAQRNKPGYY TRLP1.FRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFK.CNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAG 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKONDQRKVSQGRLAP RPPPVEKSKELAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRNQAE RERNLTEENTEKELENFKASITSSASLWHHCE		1		,			TGWGHTQYGGTGALILQKGEIRVINQTTCEN
SSVEADGRIFQAGVVSWGDGCAQRNKPGVY TRLPLFRDWIKENTGV 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTR.PSLMAGR.QHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQVEKVFGLEEYWCQCVAWS SGTITKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEFNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQFGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGFAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEGE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE]			
TRLPLFRDWIKENTGV 449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYTRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSIRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSFNLSPSASPTSSRSSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRNQAE RERNLTEENTEKELENFKASITSSASLWHHCE							
449 1799 A 3618 2 613 FVSGSPWRMDGSTERLEARRPAGRLPWSSRQ EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGGGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRNQAE RERNLTEENTEKELENFKASITSSASLWHHCE		1		[
EMTRRPSLMAGRQHGWSAQQSATVANPVPG ANPDLLPHFLGEPEDVYIVNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQGGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVPVQLSSAFRNDSYTLESRNQAE RERNLTEENTEKELENFKASITSSASLWHHCE	449	1799	A	3618	2	613	
ANPDLLPHFLGEPEDVYIVKNKPVLLVCKAV PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 A 3620 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPTAGVSCSFTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGIKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRNQAE RERNLTEENTEKELENFKASITSSASLWHHCE							
PATQIFFKCNGEWVRQVDHVIERSTDGSSGLP TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE EQGIVLPCRPPEGIPPAE MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQVPKMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE				1 1			
TMEVRINVSRQQVEKVFGLEEYWCQCVAWS SSGTTKSQKAYIRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE							
SSGTTKSQKAYTRIAYLRKNFEQEPLAKEVSL EQGIVLPCRPPEGIPPAE 450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE		1					
450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNIANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKFFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE				[İ		
450 1800 A 3620 1 2676 MEPSLGQGMDLTCPFGVSPACGAQASWSIFG ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNIANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE		f i			'		
ADAAEVPGTRGHSQQEAAMPHIPEDEEPPGE PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE	450	1800	A	3620	1	2676	
PQAAQSPAGQQGPPTAGVSCSPTPTIVLTGDA TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE							
TSPEGETDKNLANRVHSPHKRLSHRHLKVST ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE]			
ASLTSVDPAGHIIDLVNDQLPDISISEEDKKKN LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE		, ,	ŀ				
LALLEEAKLVSERFLTRRGRKSRSSPGDSPSA VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE				(l		• [
VSPNLSPSASPTSSRSNSLTVPTPPEGDEADVS SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE				1			
SPHPGEPNVPKGLADRKQNDQRKVSQGRLAP RPPPVEKSKEIAIEQKENFDPLQYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE							
RPPPVEKSKEIAIEQKENFDPLÖYPETTPKGLA PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE			•				
PVTNSSGKMALNSPQPGPVESELGKQLLKTG WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE							
WEGSPLPRSPTQDAAGVGPPASQGRGPAGEP MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE							
MGPEAGSKAELPPTVSRPPLLRGLSWDSGPEE PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE				[
PGPRLQKVLAKLPLAEEEKRFAGKAGGKLAK APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE		(i					
APGLKDFQIQVQPVRMQKLTKLREEHILMRN QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE							
QNLVGLKLPDLSEAAEQEKGLPSELSPAIEEE ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE							
ESKSGLDVMPNISDVLLRKLRVHRSLPGSAPP LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE							
LTEKEVENVFVQLSSAFRNDSYTLESRINQAE RERNLTEENTEKELENFKASITSSASLWHHCE							
RERNLTEENTEKELENFKASITSSASLWHIICE		l					
		}	· .		i		
HRETYQKLLEDIAVLHRLAARLSSRAEVVGA					1		
		L		لـــــا			HRETYQKLLEDIAVLHRLAARLSSRAEVVGA

				CB 12.2.3	Day Based and	Amino acid sequence (A=Alanine C=Cysteine,
	SEQ ID	Met	SEQ	Predicted	Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
	NO: of	hod	ID NO:	beginning	location	F=Phenylalanine, G=Glycine, H=Histidine,
	peptide		in	nucleotide	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-	l	USSN	location	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence		09/496	correspondi	acid residue	O=Glutamine, R=Arginine, S=Serine,
uence			914	ng to first		T=Threonine, V=Valine, W=Tryptophan,
- 1			ł	amino acid	of peptide	Y=Tyrosine, X=Unknown, *=Stop codon,
		ļ		residue of	sequence	/=possible nucleotide deletion, \=possible
		Ì	ł	peptide	{	nucleotide insertion
			<u> </u>	sequence		VRQEKRMSKATEVMMQYVENLKRTYEKDH
		ł	ŀ	l		AELMEFKKLANQNSSRSCGPSEDGVLRTARS
1		}		ł	ì	MSLTLGKNMPRRRVSVAVVPKFNALNLPGQ
						TPSSSSIPSLPALSESPNGKGSLPVTSALPALLE
]		}		ì	Ì	NGKTNGDPDCEASAPALTLSCLEELSQETKA
						RMEEEAYSKGFQEGLKKTKELQDLKEEEEEQ
		ļ	1	İ		KSESPEEPEEVEETEEEEKDPRSSKLEELVHFL
				1		KSESPEEPEEVEETEEEERDFRSSKLEELVIII'L
		1	1	ŀ		QVMYPKLCQHWQVIWMMAAVMLVLTVVL
		ļ				GLYNSYNSCAEQADGPLGRSTCSAAQKDSW
,		<u> </u>				WSSGLQHEQPTEQ
451	1801	Α	3623	504	198	QLIQHQTVHTGRKLYECKECGKAFNQGSTLI
ſ		ì	ì		ļ	RHORIHTGEKPYECKVCGKAFRVSSQLKQHQ
1				l		RIHTGERPYQCKELKGRGAEMLAVLAVKEQ
			1	1		NRTPVNYGK
452	1802	A	3628	2	195	MTCLHSAKAFHY*SSCSFSCEEGFALIGPEVV
				1	1	QCTALGVWTAPAPVCIAVQCQHLEALNEGT
, }					ì	MG*DYPFTAFAYGSSCKYECHTVYRVRGLD
		1				MLHSRGCYLWNGHFTT*EAISCEPLERPCH*S
						V*CSFSCEEGFALIGPEVVQCTALGVWTAPAP
l I	l	1	İ	1		VCIAVQCQHLEALNEGTMG
453	1803	A	3637	662	142	IQAKGLGIWHVPNKSPMQHWR\KGSLLRYRT
	1	}		ĺ	İ	DTGFLQTLGHNLLGIYQKYPVKYGEGKCWT
1	İ	l		ì		DNGPVIPVVYDFGDAQKTASYYSPYGQREFT
Į l		1		1	İ	AGFVQFRVFNNERAANALCAGMRVTGCNTE
1		Ì			}	HHCIGGGGYFPEASPQQCGDFSGFDWSGYGT
		1	1	1		\HVGYSSSREITE\AAVLLFYR
454	1804	A	3641	1	362	TQVHPAMLGLDELGRSGCGHCTQADLRFGD
454	100.	1		1	1	AAGRDPGQDNDRNTAEPAFPPPPRVMAAAA
	}	1		1 .		ALRAPAQSSVTFEDVAVNFSLEEWSLLNEAQ
j l	Ì					GCLYHDVMLETLTLISSLGKVLILNCDLS
455	1805	A	3646	2	414	AAAGRGASGALTGEGGGEQGRRVGLGSRAH
433	1005	1 ''	50.0] -		SLLLGPTFNSCQVSSQPPRVAGLGLPLKHEPS
		1	j	}	ļ	RPQPPSPRGPRTVRAGVPGAHPQDTPCPEFVR
	l	1				PRKVPLVGEAPGLPPEERSRGWRRDTPGLQE
		1		1	1	SRVRAPSYDDIT
456	1806	A	3656	396	8	QIVSFNSYLTLYTKNNLKSMKDLNVNTEMIK
430	1800	1^	3030	1 330	·	LLELKNIHNLG*AKFFLN*IQKALIKRKILIHW
1		1	1	1	1	P/LIKIK/SFCSLSDTIKKMKRQTIVWEQTFIIHI
1		Į.	1	1	Ì	SVKELVSRIYEAFLQFNKTVNRPVFDIKKEQK
ļ	1			1	1	l F
457	1907	+	3660	14	1961	SEAKLGGPTGMDLWQLLLTLALAGSSDAFSG
457	1807	^	3000	1 "	1	SEATAAILSRAPWSLQSVNPGLKTNSSKEPKF
1	1	ł		(TKCRSPERETFSCHWTDEVHHGTKNLGPIQLF
1		l	1		1	YTRRNTQEWTQEWKECPDYVSAGENSCYFN
	-	1	1		+	SSFTSIWIPYCIKLTSNGGTVDEKCFSVDEIVQ
]	1	1	[1	PDPPIALNWTLLNVSLTGIHADIQVRWEAPRN
	1	1	1		Į	ADIQKGWMVLEYELQYKEVNETKWKMMDP
1	1	1		1		ILTTSVPVYSLKVDKEYEVRVRSKQRNSGNY
1	1 .	1	1	[GEFSEVLYVTLPQMSQFTCEEDFYFPWLLIIF
1	1			İ		GIFGLTVMLFVFLFSKQQRIKMLILPPVPVPKI
1	Ì			1 .	Ì	KGIDPDLLKEGKLEEVNTILAIHDSYKPEFHS
Į.		1		1		DDSWVEFIELDIDEPDEKTEESDTDRLLSSDH
1		Į.		1	}	EKLHINLGVKDGDSGRTSCCEPDILETDFNAH
1	1	1	ı	1	İ	DIHEGTSEVAQPQRLKGEADLLCLDQKNQNN
3	}		3			I DIRECTNEVAUPUKLKULADLUCLDŲKNŲNN
	1		-{	1		ONTHE A CRATOORGE OF A EVALUADOR PERCAR
						SPYHDACPATOOPSVIQAEKNKPQPLPTEGAE
						SPYHDACPATQQPSVIQAEKNKPQPLPTEGAE STHOAAHIOLSNPSSLSNIDFYAQVSDITPAGS
						SPYHDACPATQQPSVIQAEKNKPQPLPTEGAE STHQAAHIQLSNPSSLSNIDFYAQVSDITPAGS VVLSPGOKNKAGMSQCDMHPEMVSLCQENF
						SPYHDACPATQQPSVIQAEKNKPQPLPTEGAE STHOAAHIOLSNPSSLSNIDFYAQVSDITPAGS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cvsteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	[09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	""		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uchec		1	/14	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i		ł	l	peptide	seducine	/=possible nucleotide deletion, \=possible
1	1					
	 	 	├──	sequence	<u> </u>	nucleotide insertion
<u> </u>		1			1	FGYQDCVTYYKAASPRTKIDAIRIPVLYLSAA
ĺ				İ	•	DDPFSTVCALPKQAAQHSPYVALLITARGGHI
	ļ			1		GFLEGLLPWQHWYMSRLLHQYAKAIFQDPE
1	1016	<u> </u>	2604			GLPDLRALLPSEDRNS
466	1816	Α	3684	3	307	SSQYIVQSKTKIFL*AAREKQ/RHTCRRFSIRLS
					(ANISSQTGEARGQWPSVFKVLKEKKLSTKKS
1	[[1	ĺ	FGQK*GR\RKTFPDKQK/LREFDTTRPTIQEML
						TGVLQG
467	1817	A	3687	2465	837	ELPTPLIAAHQLYNYVADHASSYHMKPLRMA
ł	İ	ļ				RPGGPEHNEYALVSAWHSSGSYLDSEGLRHQ
1						DDFDVSLLVCHCAAPFEEQGEAERHVLRLQF
1		!	1			FVVLTSQRELFPRLTADMRRFRKPPRLPPEPE
1			1			APGSSAGSPGEASGLILAPGPAPLFPPLAAEVG
						MARARLAQLVRLAGGHCRRDTLWKRLFLLE
{			(ł	PPGPDRLRLGGRLALAELEELLEAVHAKSIGD
						IDPQLDCFLSMTVSWYQSLIKVLLSRFPQSCR
1		1				HFQSPDLGTQYLVVLNQKFTDCFVLVFLDSH
l			1			LGKTSLTVVFREPFPVQPQDSESPPAQLVSTY
		٠ .				HHLESVINTACFTLWTRLL*GSGLDH*MSLFL
[1				ESWAYQIACQRQD*PALLGPRASQTLSDTKG
						FVTMS*GSAAPAWQQEPPSPNTHSH*PIODSR
						ESGQPRGPLGPFWGTPFGPPGRVSGVHTGWQ
						TPPRAPLPESCPLVPLTTVSHLCPLSLRVFTSHL
						DITAGHSHRDDTWVPIPALPLKHLRPPSSPFA
						LGPWVSHPLMRWVQKLSHLHSNPGTGFSMG
						GKQQRN
468	1818	A	3691	960 ·	499	QTCRKDKRAIYPHFQNE*MNEIKAI*SGTGGI
1 700	1010	Α .	3031	900	427	
						QCFHSQNDSAFFFFLFLLETEFCSAA/TVQWH
1 1				'		DFLSMQPPPPGFKQFTCLSLLSSWNYRR\PPPF
]						PGNF*FLVKTGFPHVGQTGFELLTSSDLAPLA
469	1819		2714	4242	105	SQNGGITGMSPCAWPFFFFFFGLC
409	1919	Α	3714	4747	495	MAYSWQTDPNPNESHEKQYEHQEFLFVNQP
i i						HSSSQVSLGFDQIVDEISGKIPHYESEIDENTFF
1						VPTAPKWDSTGHSLNEAHQISLNEFTSKSREL
	Į.		•			SWHQVSKAPAIGFSPSVLPKPQNTNKECSWG
i						SPIGKHHGADDSRFSILAPSFTSLDKINLEKEL
	j					ENENHNYHIGFESSIPPTNSSFSSDFMPKEENK
i i						RSGHVNIVEPSLMLLKGSLQPGMWESTWQK
						NIESIGCSIQLVEVPQSSNTSLASFCNKVKKIR
}]		J	J	ERYHAADVNFNSGKIWSTTTAFPYQLFSKTK
	ļ	l		ļ	ļ	FNIHIFIDNSTQPLHFMPCANYLVKDLIAEILH
]			ł	l	FCTNDQLLPKDHILSVWGSEEFLQNDHCLGS
					l	HKMFQKDKSVIQLHLQKSREAPGKLSRKHEE
	[- 1	[1	{	DHSQFYLNQLLEFMHIWKVSRQCLLTLIRKY
			· 1	- :		DFIILKYLLKTOENVYNIIEEVKKICSVLGCVE
	1	ļ			ŀ	TKQITDAVNELSLILQRKGENFYQSSETSAKG
		1	l	l		LIEKVTTELSTSIYQLINVYCNSFYADFQPVNV
j ,	1					PRCTSYLNPGLPSHLSFTVYAAHNIPETWVHR
	1	[ĺ	[INFPLEIKSLPRESMLTVKLFGIACATNNANLL
	Ţ	ŀ		ì		AWTCLPLFPKEKSILGSMLFSMTLQSEPPVEM
	Į.	l	1	i	ļ	ITPGVWDVSQPSPVTLQIDFPATGWEYMKPD
	ŀ	}	1			SEENRSNLEEPLKECIKHLARLSQKQTPLLLSE
]			ŀ	ļ		EKKRYLWFYRFYCNNENCSLPLVLGSAPGW
,	ì	1		j	1	
.				İ		DERTVSEMHTILRRWTFSQPLEALGLLTSSFP
		ļ		1		DQEIRKVAVQQLDNLLNDELLEYLPQLVQAV
	İ	İ	ì		ļ	KFEWNLESPLVQLLLHRSLQSIQVAHRLYWL
						LKNAENEAYFKSWYQKLLAALQFCAGKALN
]			j	j	DEFSKEQKLIKILGDIGERVKSASDHQRQEVL
	l	ļ				KKEIGRLEEFFQDVNTCHLPLNPALCIKGIDH
·					}	DACSYFTSNALPLKITFINANLMGKNISIIFKA

SEO ID SEQ ID Met SEQ Pr	redicted Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
	eginning nucleotide	D=Aspartic Acid, E=Glutamic Acid,
1.0.0.	ucleotide location	F=Phenylalanine, G=Glycine, H=Histidine,
	ocation corresponding	I=Isoleucine, K=Lysine, L=Leucine,
3000	orrespondi to last amino	M=Methionine, N=Asparagine, P=Proline,
1 200	g to first acid residue	O=Glutamine, R=Arginine, S=Serine,
	mino acid of peptide	T=Threonine, V=Valine, W=Tryptophan,
		Y=Tyrosine, X=Unknown, *=Stop codon,
		/=possible nucleotide deletion, \=possible
	eptide	
Se	equence	nucleotide insertion
		GDDLRQDMLVLQLIQVMDNIWLQEGLDMQ
	ľ	MIJYRCLSTGKDQRLVQMVPDAVTLAKIHRH
	į.	SGLIGPLKENTIKKWFSQHNHLKADYEKALR
	i	NFFYSCAGWCVVTFILGVCDRHNDNIMLTKS
	ļ	GHMFHIDFGKFLGHAQTFGGIKRDRAPFIFTS
	1	EM/EYFITEGG/KNPQHFQDFV/ELCCRAYNIIR
	ţ	KHSQLLL\NLL\EMMLYAG\LPELSGI\QDLKY
	ŀ	VYNNLRPQDTDLEATSHFTKKIKESLECFPVK
		LNNLIHTLAQMSAISPAKSTSQTFPQESCLLST
	i	TRSIERATILGFSKKSSNLYLIQVTHSNNETSL
		TEKSFEQFSKLHSQLQKQFASLTLPEFPHWW
		HLPFTNSDHRRFRDLNHYMEQILNVSHEVTN
		SDCVLSFFLSEAGQQTVEESSPVYLGEKFPDK
]		KPKVQLVISYEDVKLTILVKHMKNIHLPDGSA
	-	PSAHVEFYLLPYPSEVRRRKTKSVPKCTDPTY
1 1 1 1		NEIVVYDEVTELOGHVLMLIVKSKTVFVGAI
]	NIRLCSVPLDKEKWYPLGNSII*PLLLFSSFGM
		KSLEKDEFVGGMLLSNPIW
470 1820 A 3718 43	30 75	SHGSISILNLHQGCVFLPSLPAQGLRCYRCLA
1020 A 3/16 43] /3	VLEGASCSVVSCPFLDGVCVSQKVSV/CWQ*/
		CPWGARAEGRLSAVVDSQISCCKGDLCNAV
		VLAAGSPWALCVQLLLSLGSVFLWALL
471 1821 A 3723 89	91 494	LROSL/NSVPOAGVOWRDSSLOAPPPRFTPLS
4/1 1821 A 3/23 69	91 494	CLSLPSSWDYRRLPPCLANFLYF**RRGFTML
	·	ARMVLIS*PRDPPASASQ\STEITGGSHRAQHP TDSRDHSERSVKKSHEVISELRMKVIKCKVAF
		SKNPI
1000	43 251	<u> </u>
472 1822 A 3734 44	43 251	GFIET*NFCVSKDTSKKLS/RLPTKWKNVFAN
<u> </u>		*ISDKGLVSRICQELLRHLDAEQVSSTAGLSL
473 1823 A 3746 3	500	THASGGARSGAGWAGRGVRAGTEAGRGGIF
'		LTLSILRTRDLPSGAMSEGVDLIDIYADEEFNQ
		DPEFNNTDQIDLYDDVLTATSQPSDDRSSSTE
1 1 1 1		PPPPVRQEPSPKPNNKTPAILYTYSGLRNRRA
	İ	AVYVGSFSWWTTDQQLIQVIRSIGVYDVGEV
		KFAENRAK
474 1824 A 3753 2	5262	RPLFAREGGIYAVLVCMQEYKTSV\LVQQAG
		LAALKMLAVASSSEIPTFVTGRDSIHSLFDAQ
1 1 1 1		MTREIFASIDSATRPGSESLLLTVPAAVILMLN
		TEGCSSAARNGLLLLNLLLCNHHTLGDQIITQ
1 1 1 1	1	ELRDTLFRHSGIAPRTEPMPTTRTILMMLLNR
		YSEPPGSP\ERAALETPIIQGQDGSPELLIRSLV
		GGPSAELLLDLERVLCREGSPGGAVRPLLKRL
1 1 1 1		
		QQETQPFLLLLRTLDAPGPNKTLLLSVLRVIT
		QQETQPFLLLLRTLDAPGPNKTLLLSVLRVIT
		QQETQPFLLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE
		QQETQPFLLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI
		QQETQPFLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL
		QQETQPFLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN
		QQETQPFLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM
		QQETQPFLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKIYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG
		QQETQPFLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW
		QQETQPFLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP
		QQETQPFLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRL
		QQETQPFLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRL LQLCPRLNRVLRHEQNFADRFLPDDEAAQAL
		QQETQPFLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRL LQLCPRLNRVLRHEQNFADRFLPDDEAAQAL GKTCWEALVSPLVQNITSPDAEGVSALGWLL
		QQETQPFLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRL LQLCPRLNRVLRHEQNFADRFLPDDEAAQAL GKTCWEALVSPLVQNITSPDAEGVSALGWILD QYLEQRETSRNPLSRAASFASRVRRLCHLL
		QQETQPFLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKI'YWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRL LQLCPRLNRVLRHEQNFADRFLPDDEAAQAL GKTCWEALVSPLVQNITSPDAEGVSALGWILL DQYLEQRETSRNPLSRAASFASRVRRLCHLL VHVEPPPGPSPEPSTRPFSKNSKGRDRSPAPSP
		QQETQPFLLLRTLDAPGPNKTLLLSVLRVIT RLLDFPEAMVLPWHEVLEPCLNCLSGPSSDSE IVQELTCFLHRLASMHKDYAVVLCCLGAKEI LSKVLDKHSAQLLLGCELRDLVTECEKYAQL YSNLTSSILAGCIQMVLGQIEDHRRTHQPINIP FFDVFLRHLCQGSSVEVKEDKCWEKVEVSSN PHRASKLTDHNPKTYWESNGSTGSHYITLHM HRGVLVRQLTLLVASEDSSYMPARVVVFGG DSTSCIGTELNTVNVMPSASRVILLENLNRFW PIIQIRIKRCQQGGIDTRVRGVEVLGPKPTFWP LFREQLCRRTCLFYTIRAQAWSRDIAEDHRRL LQLCPRLNRVLRHEQNFADRFLPDDEAAQAL GKTCWEALVSPLVQNITSPDAEGVSALGWILD QYLEQRETSRNPLSRAASFASRVRRLCHLL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide]	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
				peptide		/=possible nucleotide deletion, \=possible
		1		sequence		nucleotide insertion
<u> </u>		 	 	Soquence	 	ALRSGFSGALLQQSFLTAAHMSEQFARYIDQ
		1				QIQGGLIGGAPGVEMLGQLQRHLEPIMVLSG
	ĺ	1			1	I EL ATTECHEVOLIVA A DRI L'OCCOURT DO AV
	<i>.</i>					LELATTFEHFYQHYMADRLLSFGSSWLEGAV
		Ì	1			LEQIGLCFPNRLPQLMLQSLSTSEELQRQFHLF
i					Ì	QLQRLDKLFLEQEDEEEKRL*EEEEEEEEA
	ĺ					EKELFIEDPSPAISILVLSPRCWPVSPLCYLYHP
ł	ļ	ļ	1			RKCLPTEFCDALDRFSSFYSQSQNHPVLDMG
		1	1			PHRRLQWTWLGRAELQFGKQILHVSTVQMW
1		1				LLLKFNQTEEVSVETLLKDSDLSPELLLQALV
						PLTSGNGPLTLHEGQDFPHGGVLRLHEPGPQ
						RSGEALWLIPPQAYLNVEKDEGRTLEQKRNL
						LSCLLVRILKAHGEKGLHIDQLVCLVLEAWQ
1						KGPNPPGTLGHTVAGGVACTSTDVLSCILHLL
						GQGYVKRRDDRPQILMYAAPEPMGPCRGQA
		1	1			DVPFCGSQSETSKPSPEAVATLASLQLPAGRT
						MSPQEVEGLMKQTVRQVQETLNLEPDVAQH
						LLAHSHWGAEQLLQSYSEDPEPLLLAAGLCV
1 1			ì			HQAQAVPVRPDHCPVCVSPLGCDDDLPSLCC
						MHYCCKSCWNEYLTTRIEQNLVLNCTCPIAD
ļ		·				CPAQPTGAFIRAIVSSPEVISKYEKALLRGYVE
						SCSNLTWCTNPQGCDRILCRQGLGCGTTCSK
iii			1	ı		CGWASCFNCSFPEAHYPASCGHMSQWVDDG
						GYYDGMSVEAQSKHLAKLISKRCPSCQAPIE
				İ		KNEGCLHMTCAKCNHGFCWRCLKSWKPNH
				•		KDYYNCSAMVSKAARQEKRFQDYNERCTFH
						HQAREFAVNLRNRVSAIHEVPPPRSFTFLNDA
						CQGLEQARKVLAYACVYSFYSQDAEYMDVV
	Ì					EQQTENLELHTNALQILLEETLLRCRDLASSL
						RLLRADCLSTGMELLRRIQERLLAILQHSAQD
			Ì			FRVGLQSPSVEAWEAKGPNMPGSQPQASSGP
		- 1				EAEEEEDDEDDVPEWQQDEFDEELDNDSFS
						YDESENLDQETFFFGDEEEDEDEAYD
475	1825	Α	3754	1093	96	GTSRNQHSPKTHA*RSS/WPQPPPLFLPPLQPQ
[ATGRRRRTRTQQRTAALLTDGTTKTGAAW
		i				SRRPSLCWPSRTTGAPGAK*AVLVRSATPTTN
]	İ					PPNPQSPTGAAGKLRAPGNRAG/SEPSSQEPPP
	1			l	1	DGTR\RPASITGVAQSPATRATPSLPCLHVPAP
	Į	Į		J	ļ	SRGQTLGVRTTGRASRLTVDRSRLSWPGRSA
	I		i	1		RSGGGRWRPNAPRGRWPRAP*SWEPGSWTE
			i	1		PWRWPFPAAESPPHRCIYCTNHVSPAGPARPS
	Ī	1	ŀ	ł		HVYIIRATINSISHPLCRAQSSPWEAAGVWRR
	1	Ī		ļ		PAQPAPTSDVNINLLRKPRVKRHDLIYQFLGN
<u> </u>		1	1	ļ		TLWEEGRQRPPETLQPAR
476	1826	A	3758	901	521	FFFGNGVSPCPQAGV*WHDLDSLQNLPPGFK
].						RFSYLSLPSSW\DYRHVPPRQANFCIF/M*RRG
ĺ			l	l	1	FTMLARMVSIS*PRDLPALASQSAGITGVSHH
	ļ	1				APPQMDFTFALLCFAPKGCLPRQKEGGTLNLI
477	1827	A	3761	843	575	GVISAHCNLRL/CHLPGSSNSPASASQVAGTIG
					3/3	ARTTPS*IFVFLVETGFHHVSQDGLDLLNFVI
	j			i	Ì	RPRRPLKVLGLOACTRARLPSPLKEL
478	1828	A	3763	267	1240	
770		^	3703	20/	1240	HLLSFHLWSASLDCLEQLSQERHVKGMLLGP
		j		ļ		PPVNESTKPSPSPWKLTPPMCSIPPVFPPKSGS
		.]	İ	Į.	İ	PTTSWS/PSGHSKLEVERAQTGPFCLHIYCP*P
		- 1	1	1	j	GVTDNTTSLLHYIPFPRL\SGLVCFPAH*FPSY
- 1		f		ł		WTGHSFASQAWLRQVPEVSKHLQCPSAESLL
		ļ		1		TMEYHQPEDPAPGKAGTAEAVIPENHEVLAG
	ļ	1		l	1	PDEHPQDTDARDADGEAREREP/RRPSFAA*P
- 1	1	- 1	i	i	İ	VWGQP\ESPLPEASSAPPGPTLGTLPEVETIRA
			ļ	,	j	CSMPQELP*SPRTRQPEPDFYCVKWIPWKGE
i	1					QTPIITQSTNGPLPSPCHHEHPLSSVEGEAPPA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
479	1829	A	3766	2	2152	EGSDHIG YSPIRLLEVCVPLPKIFIKRQAPLKVSLLQDLK DFFQKVSQVYVAIDERLASLKTDTFSKTREEK MEDIFAQKEMEEGEFKNWIEKMQARLMSSS VDTPQQLQSVFESLIAKKQSLCEVLQAWNNR LQDLFQQEKGRKRPSVPPSFGRLRQGEESKIS AMDASPRNISPGLQNGEKEDRFLTTLSSQSST SSTHLQLPTPPEVMSEQSVGGPPELDTASSSE DVFDGHLLGSTDSQVKEKSTMKAIFANLLPG NSYNPIPFFFDPDKHYLMYEHERVPIAVCEKE PSSIIAFALSCKEYRNALEELSKATQWNSAEE GLPTNSTSDSRPKSSSPIRLPEMSGGQTNRTTE TEPQPTKKASGMLSFFRGTAGKSPDLSSQKRE TLRGADSAYYQVGQTGKEGTENQGVEPQDE VDGGDTQKKQLINPHVELQFSDANAKFYCRL YYAGEFHKMREVILDSSEEDFIRSLSHSSPWQ ARGGKSGAAFYATEDDRPILKQMPRLEVQSF LDFAPHYFNYITNAVQQKRPTALAKILGVYRI GYKNSQNTEKKLDLLVMENLFYGRKMAQ VFDLKGSLRNRNVKTDTGKESCDVVLLDENL LKMVRDNPLYIRSHSKAVLRTSIHSDSHFLSS HLIIDYSLLVGRDDTSNELVVGIIDYIRTFTWD KKLEMVVKSTGILGGQG*MPTVVSPELYRTR
480	1830	A	3777	251	3	FCEAMDNYFLMVPDHCTGLGLNC QGCGSAGTLIHY**ECKMVQLLWKTV*QFLI KLNIKDPAITLDVYPNEVKNYVRTKTYTQMF I/ANFIMAKSWKQPTHPSVRT
481	1831	Α.	3779	333	3	EAAIRQPEPNILDVNQIFKDLAMIIHDQGDLID SIEANAESSEVLVERAPGQLQRPA\YYQKKSR KKMCLVVLVQTAIILICERIM*VVYTTKWSPPI VLPVSCFQGQKFN
482	1832	Α	3780	2	371	TGGRQGKNDHTSITEKPSRDFNRHLITQNI*M PNQDMKSSSNSLIIRKVQIKPTILYHHIFTRKA KMKTTDKTKYR*GFKAITTLIHCSQDCKLQ*S /L*ENHFMIFPKAEQHITYDTTIPFLR
483	1833	A	3787	43	448	LMKDLSPYVMETHYILNRLNER/RSMWRHIIG KLPNTKDQEKILKAIRGRREVIQGS/RQQYRR PAAFSAAEKARRLWCS/VFNIERRNL/CEYPTK LSFNIKGEMTFSDKTEFTTNRPSLKMLLKDRI QEEGKMF*KEKCFKRKE
484	1834	A	3798	1	727	FFFFETESRSVAQAGVQWCNLGSLQALPPGF\ SHSPASASRVAGTTGTRH*ARLIFYIFSRDGVS PC*PGWS*SPDLVIRPP\RLPKCWDYRREPPRP A*FFVFLVE\QGFTMLARMVSIS*PQ\CDL\PAS VSQNAGITGVSHCAWPCLHFCFFGFFFEMESC SVAQAEVQWHDLRSLQAPPPGFTFFSCLSLPG SWDYRRPPPRPANF\CIFSRDGVSPC*PGWSRS PDLVIRPPRPPKVLGLQA
485	1835	Α	3802	1	239	FFFFEMECLTVSQAGVQWYNLHSLQPLPPGF KQFSC\LSLPSSWD*RVPTSRPAKF/CVIF*DGV SHCQPGWSAVVQPPLH
486	1836	A	3811	378	98	RYD*SSQSENIP\QKEFLLKYP*CTATLGMRN MSIMKKKSIFSAEFYKVSLPSLLL\HLLAIEWG FHIEIQLTIHQHFLNYELESDFVHIVEYM
487	1837	A	3814	771		FDPDWTRAAGIRHEKKPKALAYRRENSPGDL PPPPLPPPEEEASWAL/GAEGSRQHVLPGAGA QWGEESGPGRAPGSPAGAPPR*RGLAP\NSRP SFLSRGQGTSTCSTAGSNSSRGSSSSRGSRGPG RSRSRSQSRSQSQRPGQKRREEPR

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	Í	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	1	914	amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Scrine, T=Threonine, V=Valine, W=Tryptophan,
		i		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	İ			peptide	Sequence	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
488	1838	A	3818	1	781	FRACLLELIPYAPTLSWTACPPAMAGPRGLLP
		ı		Ì	1	LCLLAFCLAGFSFVRGQVLFKGCDVKTTFVT
						HVPCTSCAAIKKQTCPSGWLRELPDQITQDCR
				i		YEVQLGGSMVSMSGCRRKCRKQVVQKACCP
	j .	J				GYWGSRCHECPGGAETPCNGHGTCLDGMDR
		ĺ				NGTCVCQENFRGSACQECQDPNRFGPDCQSV
						CSCVHGVCNHGPRGDGSCLCFAGYTGPHCD
						QELPVWQELGFPQNNPRLRKAPNCKCLPG*H RNGLIATPNPCRP
489	1839	Α	3822	934	669	FFFSEMESRSVTRLECSGAISAHLRLLGSSNSP
	1007	* -	3022	334	007	ASAS*VAGTIGACHHAQLIFVFLVETGFHHVG
		ĺ			l	QDGLDLL/NLMIHPPRPPKVLGFQA
490	1840	A	3825	79	9748	GCQSCWPAWPRLRRRGPASAGARLGRKAPW
			ŀ			GLPGRVQDGRPLRFCFYLRPRAPFIAPVLSGA
						ASRPEASGDCRAGRETAMATLEKLMKAFESL
						KSFQQQQQQQQQQQQQQQQQQPPPP
						PPPPPPQLPQPPPQAQPLLPQPQPPPPPPPPPPPPPPPP
						GPAVAEEPLHRPKKELSATKKDRVNHCLTIC
]]		ENIVAQSVRNSPEFQKLLGIAMELFLLCSDDA
						ESDVRMVADECLNKVIKALMDSNLPRLQLEL YKEIKKNGAPRSLRAALWRFAELAHLVRPOK
						CRPYLVNLLPCLTRTSKRPEESVQETLAAAVP
						KIMASFGNFANDNEIKVLLKAFIANLKSSSPTI
						RRTAAGSAVSICQHSRRTQYFYSWLLNVLLG
						LLVPVEDEHSTLLILGVLLTLRYLVPLLQQQV
						KDTSLKGSFGVTRKEMEVSPSAEQLVQVYEL
				,		TLHHTQHQDHNVVTGALELLQQLFRTPPPEL
1						LQTLTAVGGIGQLTAAKEESGGRSRSGSIVELI
	1				i	AGGGSSCSPVLSRKQKGKVLLGEEEALEDDS ESRSDVSSSALTASVKDEISGELAASSGVSTPG
1	1					SAGHDITEQPRSQHTLQADSVDLASCDLTSS
						ATDGDEEDILSHSSSQVSAVPSDPAMDLNDG
						TQASSPISDSSQTTTEGPDSAVTPSDSSEIVLD
						GTDNQYLGLQIGQPQDEDEEATGILPDEASEA
- 1]			FRNSSMALQQAHLLKNMSHCRQPSDSSVDKF
ĺ	l	i				VLRDEATEPGDQENKPCRIKGDIGQSTDDDS
ļ			j			APLVHCVRLLSASFLLTGGKNVLVPDRDVRV
1	ļ		1			SVKALALSCVGAAVALHPESFFSKLYKVPLD
- 1	1	i	ĺ	1	·	TTEYPEEQYVSDILNYIDHGDPQVRGATAILC GTLICSILSRSRPHVGDWMGTIRTLTGNTFSI.
				i		ADCIPLLRKTLKDESSVTCKLACTAVRNCVM
	- 1					SLCSSSYSELGLQLIIDVLTLRNSSYWLVRTEL
	}					LETLAEIDFRLVSFLEAKAENLHRGAHHYTGL
		1		1		LKLQERVLNNVVIHLLGDEDPRVRHVAAASL
				İ		IRLVPKLFYKCDQGQADPVVAVARDQSSVYL
		Ì	1			KLLMHETQPPSHFSVSTITRIYRGYNLLPSITD
			1			VTMENNLSRVIAAVSHELITSTTRALTFGCCE
		ļ		i	l	ALCLLSTAFPVCIWSLGWHCGVPPLSASDESR
- 1	1		1	ŀ	- 1	KSCTVGMATMILTLLSSAWFPLDLSAHQDAL
1		l	- 1			ILAGNLLAASAPKSLRSSWASEEEANPAATK
			l		l	QEEVWPALGDRALVPMVEQLFSHLLKVINIC
		- 1	ŀ		l	AHVLDDVAPGPAIKAALPSLTNPPSLSPIRRK GKEKEPGEOASVPLSPKKGSEASAASROSDTS
ŀ	i	1	i		Í	GPVTTSKSSSLGSFYHLPSYLKLHDVLKATHA
		ĺ			ļ	NYKVTLDLQNSTEKFGGFLRSALDVLSOILEL
ł			1		į	ATLQDIGKCVEEILGYLKSCFSREPMMATVC
ŀ			1		ļ	VQQLLKTLFGTNLASOFDGLSSNPSKSOGRA
	ļ			ŀ	ļ	QRLGSSSVRPGLYHYCFMAPYTHFTQALADA
ł	- 1	- 1	- 1	1	ł	SLRNMVQAEQENDTSGWFDVLQKVSTQLKT
i i	- }					NLTSVTKNRADKNAIHNHIRLFEPLVIKALKO

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first	to last amino	M=Methionine, N=Asparagine, P=Proline,
I didice			714	amino acid	acid residue of peptide	Q=Ghutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			1	peptide	Sedantice	/=possible nucleotide deletion, \=possible
İ		ļ		sequence	1	nucleotide insertion
						YTTTTCVQLQKQVLDLLAQLVQLRVNYCLL
1		ļ ,	ļ		1	DSDQVFIGFVLKQFEYIEVGQFRESEAIIPNIFF
i						FLVLLSYERYHSKQIIGIPKIIQLCDGIMASGR
						KAVTHAIPALQPIVHDLFVLRGTNKADAGKE
						LETQKEVVVSMLLRLIQYHQVLEMFILVLQ0
Ì						CHKENEDKWKRLSRQIADIILPMLAKQQMHI
Į.						DSHEALGVLNTLFEILAPSSLRPVDMLLRSMF
						VTPNTMASVSTVQLWISGILAILRVLISQSTED
						IVLSRIQELSFSPYLISCTVINRLRDGDSTSTLE EHSEGKQIKNLPEETFSRFLLQLVGILLEDIVT
	ļ					KQLKVEMSEQQHTFYCQELGTLLMCLIHIFKS
	i					GMFRRITAAATRLFRSDGCGGSFYTLDSLNLR
1 1			ł			ARSMITTHPALVLLWCQILLLVNHTDYRWW
						AEVQQTPKRHSLSSTKLLSPQMSGEEEDSDLA
					-	AKLGMCNREIVRRGALILFCDYVCQNLHDSE
						HLTWLIVNHIQDLISLSHEPPVQDFISAVHRNS
						AASGLFIQAIQSRCENLSTPTMLKKTLQCLEGI
1 1			•			HLSQSGAVLTLYVDRLLCTPFRVLARMVDIL
						ACRRVEMLLAANLQSSMAQLPMEELNRIQEY LQSSGLAQRHQRLYSLLDRFRLSTMQDSLSPS
			Ī			PPVSSHPLDGDGHVSLETVSPDKDWYVHLVK
		- 1	,	j	j	SQCWTRSDSALLEGAELVNRIPAEDMNAFM
		- 1		i		MNSEFNLSLLAPCLSLGMSEISGGQKSALFEA
]	-	1			AREVILARVSGTVQQLPAVHHVFQPELPAEP
				•		AAYWSKLNDLFGDAALYQSLPTLARALAQY
			1			LVVVSKLPSHLHLPPEKEKDIVKFVVATLEAL
		l	ŀ	<i>'</i>		SWHLIHEQIPLSLDLQAGLDCCCLALQLPGL
	l	- 1	ĺ			WSVVSSTEFVTHACSLIYCVHFILEAVAVQPG
1		1		•		EQLLSPERRTNTPKAISEEEEEVDPNTQNPKYI TAACEMVAEMVESLQSVLALGHKRNSGVPA
		ı				FLTPLLRNIIISLARLPLVNSYTRVPPLVWKLG
	Į.	- 1	i			WSPKPGGDFGTAFPEIPVEFLQEKEVFKEFIYR
1		- 1			ļ	INTLGWTSRTQFEETWATLLGVLVTQPLVME
						QEESPPEEDTERTQINVLAVQAITSLVLSAMT
	1			I		VPVAGNPAVSCLEQQPRNKPLKALDTRFGRK
			1	-		LSIIRGIVEQEIQAMVSKRENIATHHLYQAWD
1						PVPSLSPATTGALISHEKLLLQINPERELGSMS
[ĺ	i	ĺ		•	YKLGQVSIHSVWLGNSITPLREEEWDEEEEEE
				l		ADAPAPSSPPTSPVNSRKHRAGVDIHSCSQFL LELYSRWILPSSSARRTPAILISEVVRSLLVVS
ļ I			.	!		DLFTERNQFELMYVTLTELRRVHPSEDEILAQ
			ŀ			YLVPATCKAAAVLGMDKAVAEPVSRLLESTL
j j			j	İ	ļ	RSSHLPSRVGALHGVLYVLECDLLDDTAKOL
		-]	[ſ	IPVISDYLLSNLKGIAHCVNIHSQQHVLVMCA
		†	İ	* "	İ	TAFYLIENYPLDVGPEFSASIIOMCGVMLSGS
	1	İ				EESTPSIIYHCALRGLERLLLSEQLSRLDAESL
1 1	- 1	i	ł	ł	ł	VKLSVDRVNVHSPHRAMAALGLMLTCMYT GVEKVSBGBTSDBNBAARDSESURVAN FRANCE
.						GKEKVSPGRTSDPNPAAPDSESVIVAMERVS VLFDRIRKGFPCEARVVARILPQFLDDFFPPQ
						DIMNKVIGEFLSNQQPYPQFMATVVYKVFOT
	ļ	- 1		İ		LHSTGQSSMVRDWVMLSLSNFTQRAPVAMA
	1	- 1		ł		TWSLSCFFVSASTSPWVAAILPHVISRMGKLE
[[1		- 1		ł	QVDVNLFCLVATDFYRHQIEEELDRRAFQSV
407						LEVVAAPGSPYHRLLTCLRNVHKVTTC
491	1841	A	3826	469	302	SNPPASASRVAGITGVHQHAWLIFVFLVEMEF
402	1042	, 	2026	300		HHVGQAVLKLLISGDLPVSASQSA
492	1842	A	3836	392	88	VAPSPMIMPDLYFYRDPEEEKEE*AAAEK\EE
	ĺ		1		1	FQSEWTAVV/P/EFTATQSEVADWFKDMQVP
		- 1	1	1	j	SVPIQQFPTEDWST*PIMNDWSATSTAQTTE WVRITTEWP
<u> </u>				L		WANTIEML

SEQ ID	SEO ID	Mat	CEO	Danding	Dunding and	I A - t - c - t - c - c - c - c - c - c - c
NO: of	NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in No.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	{	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
l		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		l		peptide	304-01100	/=possible nucleotide deletion, \=possible
İ		1	1	sequence		nucleotide insertion
493	1843	A	3838	19	380	TPSDMNRAFETDTQSIGEKNRSPSEPDYFERK
						KFKRS*EKAHIRYKIDOPEDIPLK\EFLCKHSK
1		1				CTATLSMRNMSLMKKKCSFSEEF\LAFFPSLL
l	İ	l	1	}	1	VCHLLAIKLGFYIEIHLTTFNNTF
494	1844	A	3845	2	352	FFFLRRSL/DSVAQAEAQWL/ELGLLQAPPPGF
}	ł	i	1			KPISLP\GLPSSWDYGRPPPCPANFCIF/M*RRG
						FTVLARMVLIS*PCDPPTLASQGTAITGMSYH
1	İ	}				ARPODIDFLYAHOGRCWFRLL
495	1845	Α	3847	1774	40	DIFFRRAKEGMGQDEAQFSVEMPLTGKAYL
		l			1	WADKYRPRKPRFFNRVHTGFEWNKYNOTHY
1	ļ	1	j		ļ	DFDNPPPKIVQGYKFNIFYPDLIDKRSTPEYFL
		1				EACADNKDFAILRFHAGPPYEDIAFKIVNREW
1		l			İ	EYSHRHGFRCQFANGIFQLWFHFKRYRYRR*
			1			RPWGTAGRCPRGHSKGASVKLVVTPGPLSGL
	1	ĺ	1		Į	QGRGFTSHLRPHLSFARPQFPPI*KGGHH*AC
]	ļ]		ļ	HGELRRHWDRLA*GPDATEGALGASFEHEG
	İ					GQQPPADLTVQADTLHRPSARLGGAHRACPK
		}				RRPHRVLWRWARGAWAWRCQAREKQETQG
						QPCHITGHPLGREAEPAAAGAAPALAHRPPF
			1		ļ	ARTGSTE\PGPCWRPIRHCRRDPLWTPTLC\RD
	,					WPPTHPVLAGGVHFPAAG/IGGCVEVPVSVN
		ļ				VMGTKSH*AVLPPPPSTGPGGQGLPEGWGLE
		Ì				KGEGLPPGIPPPGLLTGPW\SMRPVTPSFAHIR
i 1						TVAPSHSPFSGQEGRGPHGCHSPGR\SGP\AGR
						LVLQHPTGTSPTEAKRKVPPGPPEGHPTSPVT
						SPRPPTAPPRHPASSGNSSVCFSKKTCRWEKK
				<u> </u>		SFVLMELAYWQDRMFF
496	1846	Α	3849	830	442	AKSPLPLG*IQWR/NLGSLKLRLPGFK*FTCLG
			i i	•		LLSSWDYRSLPPRPVNFCILVELGFHHVDQAG
						LKLLTSSALPALASQSAEITGMSHRIWPLPLLR
						RPPVIRIRAPPQRLPFNLITSLKALSPNMATF
497	1847	Α	3859	2	393	ALRKTRRDGIARTGAQPAASWKGTNNYPWR
						LEMAGRPGSQEQSKDRGTGSLPPPSQRPLGPS
			}			PEGAGPSPPPPGIPRGGGSSSSEGP/PQLLFVPR
				-		RFPAPKKGLPSDTPHSKAPPTPHLILGGEDSQ
100	0070	<u></u>				VPIL
498	1848	Α	3860	253	634	KNASTVYSSQGDPKSFFFLLRWSLALVAQAG
						EQ*RDLSSLQPPPPGFK*FSCLSLPSSWD\YRCP
		ĺ				LPCLANF*FLVETGFHHVGQADLKLLTSGDP
400	1040		12062	400	2/2	PTSASESAGITGVSHRAWPRIHFLYWKTFFL
499	1849	Α	3863	423	263	APSQISVAFLYAA/DKLFEKEI*KKIPFIIAS/DKI
						KIGINLTKEVKYLYTENYITLMKEIK/DTDKW
			-			KDILY*WIGKINI*KMSTPPKAIYRFNAIPTKIP
			•		- 4	MTFFTEIEKSIIKFIWNIIKKPPNTQSNIEQKE*S
						FCSILLWVFGGFLWFHMNFMIDFSISVKNVIGI
500	1960		2005		15045	LVGIALNL
טטכ	1850	A	3865	2	15246	LPRGCLWCLQRSPIPARPQPSRPARSPLPLFP
]]			DLRPWASDLDIMGDAEGEDEVQFLRTDDEV
					l	VLQCSATVLKEQLKLCLAAEGFGNRLCFLEP
					l	TSNAQNVPPDLAICCFVLEQSLSVRALQEML
				1	ļ	ANTVEAGVESSQGGGHRTLLYGHAILLRHAH
			[[i	SRMYLSCLTTSRSMTDKLAFDVGLQEDATGE
					ł	ACWWTMHPASKQRSEGEKVRVGDDIILVSVS
	. [[[ſ	SERVLHLSTASGELQVDASFMQTLWNMNPIC
				1	ļ	SRCEEGFVTGGHVLRLFHGHMDECLTISPADS
	[ſ	ſ	DDQRRLVYYEGGAVCTHARSLWRLEPLRIS
					ļ	WSGSHLRWGQPLRVRHVTTGQYLALTEDQG
						LVVVDASKAHTKATSFCFRISKEKLDVAPKR
	l			1		DVEGMGPPEIKYGESLCFVQHVASGLWLTYA

SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
1 .				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						APDPKALRLGVLKKKAMLHQEGHMDDALSL
1						TRCQQEESQAARMIHSTNGLYNQFIKSLDSFS
						GKPRGSGPPAGTALPIEGVILSLQDLIIYFEPPS
1						EDLQHEEKQSKLRSLRNRQSLFQEEGMLSMV
1	,					LNCIDRLNVYTTAAHFAEFAGEEAAESWKEI
J						VNLLYELLASLIRGNRSNCALFSTNLDWLVS
						KLDRLEASSGILEVLYCVLIESPEVLNIIQENHI
						KSIISLLDKHGRNHKVLDVLCSLCVCNGVAV
l l		1				RSNQDLITENLLPGRELLLQTNLINYVTSIRPN
1 1			1			IFVGRAEGTTQYSKWYFEVMVDEVTPFLTAQ
						ATHLRVGWALTEGYTPYPGAGEGWGGNGV
1						GDDLYSYGFDGLHLWTGHVARPVTSPGQHL
1					İ	LAPEDVISCCLDLSVPSISFRINGCPVQGVFESF
		· [1		NLDGLFFPVVSFSAGVKVRFLLGGRHGEFKF
						LPPPGYAPCHEAVLPRERLHLEPIKEYRREGP
1 1						RGPHLVGPSRCLSHTDFVPCPVDTVQIVLPPH
						LERIREKLAENIHELWALTRIEQGWTYGPVRD
1				1		DNKRLHPCLVDFHSLPEPERNYNLOMSGETL
1			ĺ	ĺ		KTLLALGCHVGMADEKAEDNLKKTKLPKTY
						MMSNGYKPAPLDLSHVRLTPAQTTLVDRLAE
						NGHNVWARDRVGQGWSYSAVQDIPARRNPR
}			- 1		,	LVPYRLLDEATKRSNRDSLCQAVRTLLGYGY
						NIEPPDQEPSQVENQSRCDRVRIFRAEKSYTV
1						QSGRWYFEFEAVTTGEMRVGWARPELRPDV
1 1			1			ELGADELAYVFNGHRGQRWHLGSEPFGRPW
					·	QPGDVVGCMIDLTENTIIFTLNGEVLMSDSGS
			i			ETAFREIEIGDGFLPVCSLGPGQVGHLNLGQD
1			•			VSSLRFFAICGLQEGFEPFAINMORPVTTWFS
	- 1	- 1				KGLPQFEPVPLEHPHYEVSRVDGTVDTPPCLR
	1				Ì	LTHRTWGSQNSLVEMLFLRLSLPVQFHQHFR
	ĺ	ļ				CTAGATPLAPPGLQPPAEDEARAAEPDPDYE
		1				NLRRSAGGWSEAENGKEGTAKEGAPGGTPQ
1 1		1			į.	AGGEAQPARAENEKDATTEKNKKRGFLFKA-
1 (- 1	1	i]	KKVAMMTQPPATPTLPRLPHDVVPADNRDD
1		i				PEILINTTTYYYSVRVFAGQEPSCVWAGWVT
1 1		- 1	ŀ		}	PDYHQHDMSFDLSKVRVVTVTMGDEQGNV
1 1			l	!		HSSLKCSNCYMVWGGDFVSPGQQGRISHTDL
		l	l			VIGCLVDLATGLMTFTANGKESNTFFQVEPN
1		l		1		TKLFPAVFVLPTHQNVIQFELGKQKNIMPLSA
	ļ	l	l			AMFQSERKNPAPQCPPRLEMQMLMPVSWSR
	Ì	ļ	l	1	j	MPNHFLQVETRRAGERLGWAVQCQEPLTMM
		1	[ļ		ALHIPEENRCMDILELSERLDLQRFHSHTLRL
	ļ	- 1		i		YRAVCALGNNRVAHALCSHVDQAQLLHALE
j l		1				DAHLPGPI.RAGYYDLLISIHI.ESACRSRRSML
ļ ([SEYIVPLTPETRAITLFPPGRSTENGHPRHGLP
]	i	ļ		l	. 1	GVGVTTSLRPPHHFSPPCFVAALPAAGAAEAP
	ļ	l	ł	l		ARLSPAIPLEALRDKALRMLGEAVRDGGQHA
1 1		J	J	ì	J	RDPVGASVEFQFVPVLKLVSTLLVMGIFGDE
]			ļ	i	ļ	DVKQILKMIEPEVFTEEEEEEDEEEGEEEDEE
		Į	i	j	1	EKEEDEEETAQEKEDEEKEEEEAAEGEKEEG
[ł	i]			LEEGLLQMKLPESVKLQMCHLLEYFCDQELQ
	i	ľ	1	ľ	ľ	HRVESLAAFAERYVDKLQANQRSRYGLLIKA
		- 1			1	FSMTAAETARRTREFRSPPQEQINMLLQFKDG
	-	- 1	1		1	TDEEDCPLPEEIRQDLLDFHQDLLAHCGIQLD
}	ļ	- 1	}	J	1	GEEEPEETTLGSRLMSLLEKVRLVKKKEEK
		1		ļ	1	PEEERSAEESKPRSLQELVSHMVVRWAQEDF
	ĺ	- 1	İ	Ī	l	VQSPELVRAMFSLLHRQYDGLGELLRALPRA
	ł	- 1		ĺ	ł	YTISPSSVEDTMSLLECLGQIRSLLIVQMGPQE
1 1	l	l	1	ł	f	ENLMIQSIGNIMNNKVFYQHPNLMRALGMHE
L	1					TVMEVMVNVLGGGESKEIRFPKMVTSCCRFL

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	Ì	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
ì		l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ì	1	l		peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1						CYFCRISRQNQRSMFDHLSYLLENSGIGLGM
	ļ					QGSTPLDVAAASVIDNNELALALQEQDLEKV
						VSYLAGCGLQSCPMLVAKGYPDIGWKPCGG
	[ERYLDFLRFAVFVNGESVEENANVVVRLLIR
						KPECFGPALRGEGGSGLLAAIEEAIRISEDPAR
						DGPGIRRDRRREHFGEEPPEENRVHLGHAIMS
						FYAALIDLLGRCAPEMHLIQAGKGEALRIRAI
]						LRSLVPLEDLVGIISLPLQIPTLGKDGALVQPK
			1	•		MSASFVPDHKASMVLFLDRVYGIENQDFLLH VLDVGFLPDMRAAASLDTATFSTTEMALAV
						NRYLCLAVLPLITKCAPLFAGTEHRAIMVDS
						MLHTVYRLSRGRSLTKAQRDVIEDCLMSLCR
						YIRPSMLQHLLRRLVFDVPILNEFAKMPLKLL
						TNHYERCWKYYCLPTGWANFGVTSEEELHL
						TRKLFWGIFDSLAHKKYDPELYRMAMPCLC
						AIAGALPPDYVDASYSSKAEKKATVDAEGNF
						DPRPVETLNVIIPEKLDSFINKFAEYTHEKWAF
		•				DKIQNNWSYGENIDEELKTHPMLRPYKTFSE
						KDKEIYRWPIKESLKAMIAWEWTIEKAREGE
						EEKTEKKKTAKISQSAQTYDPREGYNPQPPDL
						SAVTLSRELQAMAEQLAENYHNTWGRKKKQ
						ELEAKGGGTHPLLVPYDTLTAKEKARDREKA
				Ì		QELLKFLQMNGYAVTRGLKDMELDSSSIEKR
						FAFGFLQQLLRWMDISQEFIAHLEAVVSSGRV
1						EKSPHEQEIKFFAKILLPLINQYFTNHCLYFLS
1			İ			TPAKVLGSGGHASNKEKEMITSLFCKLAALV
				. '		RHRVSLFGTDAPAVVNCLHILARSLDARTVM
				•		KSGPEIVKAGLRSFFESASEDIEKMVENLRLG
		1		•		KVSQARTQVKGVGQNLTYTTVALLPVLTTLF
		- 1	}	}	ł	QHIAQHQFGDDVILDDVQVSCYRTLCSIYSLG
1 1	.]	1		j		TTKNTYVEKLRPALGECLARLAAAMPVAFLE PQLNEYNACSVYTTKSPRERAILGLPNSVEEM
			i		1	CPDIPVLERLMADIGGLAESGARYTEMPHVIE
						ITLPMLCSYLPRWWERGPEAPPSALPAGAPPP
						CTAVTSDHLNSLLGNILRIIVNNLGIDEASWM
		1	ľ			KRLAVFAQPIVSRARPELLQSHFIPTIGRLRKR
		i				AGKVVSEEEQLALEAKAEAQEGELLVRDEFS
			- 1			VLCRDLYALYPLLIRYVDNNRAQWLTEPNPS
]	1	ļ	- 1			AEELFRMVGEIFIYWSKSHNFKREEQNFVVQ
		İ	1	ļ		NEINNMSFLTADNKSKMAKAGDIQSGGSDQE
	ļ		l	ļ		RTKKKRRGDRYSVQTSLIVATLKKMLPIGLN
1	. 1	ļ	ł	!	}	MCAPTDQDLITLAKTRYALKDTDEEVREFLH
	.	. 1	j	İ	į	NNLHLQGKVEGSPSLRWQMALYRGVPGREE
		·	_ 1			DADDPEKIVRRVQEVSAVLYYLDQTEHPYKS
1 1		Ì				KKAVWHKLLSKQRRRAVVACFRMTPLYNLP
	1	ŀ	- 1	ŀ		THRACNMFLESYKAAWILTEDHSFEDRMIDD
1	· 1	i	j	Ĭ	ľ	LSKAGEQEEEEEEVEEKKPDPLHQLVLHFSRT
		ł		ļ		ALTEKSKLDEDYLYMAYADIMAKSCHLEEG
1	1	1		1		GENGEAEEVEVSFEEKQMEKQRLLYQQARL
1	. !	į		1	1	HTRGAAEMVLQMISACKGETGAMVSSTLKL
ļ ļ	·]	j	j	į	ļ	GISILNGGNAEVQQKMLDYLKDKKEVGFFQS
		i		i	1	IQALMQTCSVLDLNAFERQNKAEGLGMVNE DGTVINRQNGEKVMADDEFTODLFRFLOLLC
	ł	İ		ŀ		EGHNNDFONYLRTQTGNTTTINIICTVDYLL
				l	1	RLQESISDFYWYYSGKDVIEEQGKRNFSKAM
				j	1	SVAKQVFNSLTEYIQGPCTGNQQSLAHSRLW
	ļ	J	J	ļ	ļ	DAVVGFLHVFAHMMMKLAQDSSQIELLKEL
	ŀ	.	1			LDLQKDMVVMLLSLLEGNVVNGMIAROMV
				l		DMLVESSSNVEMILKFFDMFLKLKDIVGSEAF
L. I	ł	-	- 1	l		QDYVTDPRGLISKKDFQKAMDSQKQFSGPEI
						, , , , , , , , , , , , , , , , , , , ,

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *Stop codon, /-possible nucleotide deletion, \-possible nucleotide insertion
						QFLLSCSEADENEMINCEEFANRFQEPARDIG FNVAVLLTNLSEHVPHDPRLHNFLELAESILE YFRPYLGRIEIMGASRRIERIYFEISETNRAQW EMPQVKESKRQFIFDVVNEGGEAEKMELFVS FCEDTIFEMQIAAQISEPEGFPETDEDEGAGA AEAGAEGAEEGAAGLEGTAATAAAGATARV VAAAGRALRGLSYRSLRRVRRLRRLTAREA ATAVAALLWAAVTRAGAAGAGAAAGALGL LWGSLFGGGLVEGAKKVTVTELLAGMPDPT SDEVHGEQPAGPGGDADGEGASEGAGDAAE GAGDEEAVHEAGPGGADGAVAVTDGGPFR PEGAGGLGDMGDTTPAEPPTPEGSPILKRLG VDGVEEELPPEPEPEPELEPEKADAENGEK EEVPEPTPEPPKKQAPPSPPPKKEEAGGEFWG ELEVQRVKFLNYLSRNFYTLRFLALFLAFAIN FILLFYKVSDSPPGEDDMEGSAAGDVSGAGS GGSSGWGLGAGEEAEGDEDENMVYYFLEES TGYMEPALRCLSLLHTLVAFLCIIGYNCLKVP LVIFKREKELARKLEFDGLYTTEQPEDDDVKG QWDRLVLNTPSFPSNYWDKFVKRKVLDKHG DIYGREIIAELLGMDLATLEITAHNERKPNPP PGLLTWLMSIDVKYQIWKFGVIFTDNSFLYLG WYMVMSLLGHYNNFFFAAHLLDIAMGVKTL
	·					RTILSSVTHNGKQLVMTVGLLAVVVYLYTVV AFNFFRKFYNKSEDEDEPDMKCDDMMTCYL FHMYVGVRAGGGIGDEIEDPAGDEYELYRVV FDITFFFFVIVILLAIIQGLIIDAFGELRDQQEQV KEDMETKCFICGIGSDYFDTTPHGFETHTLEE HNLANYMFFLMYLINKDETEHTGQESYVWK MYQERCWDFFPAGDCFRKQYEDQLS
501	1851	A	3869	467	665	VIVAIYCQLIFDKGAKTIQ*PFQQIAL/CKRMK LGPCFTPCGKINSEWIRELSVRVKTIKHLEIGV N
502	1852	A	3888	1042	724	SGMQWRDLTPLQPLPPRFKQFSCLSLPGSWD YRHAP\PLLTNF*FLVEMGFCYVGQAGRKLL ASSDQSALASQSAGITGISTAPGPPFFFLNFEA GSCSVAQAGVQ
503	1853	A	3891	1773	1193	EVDSQSGVQ*QAPGSLQLQTPGLK/VSCLLSR QDYRSSLPHLASCCYYYYYY/VFL*RRGLTTL VQGGLKLLPSSNPFASAP*TAGITGMSHCAGP HFNF*MFRKISCIRE*F*HTRIYDIPFLILFFKET WVLLCYPGWPQIPGLKPSSCLRLLSSWDHRC APPCPASFFIFHVDRVSPPCPGLVSITFKMLLL L
504	1854	В	3896	279	70	MVSKSKSILMSYNHVELITSDMKKMPEAFRR TQKHTTYLIPYQVIFWSTGKDAMRSFMMPFY OKEYYENO*
505	1855	A	3899	2	1396	EPGYPTKKTWFDKPDFNRTNSPGFQKKVQFG NENTKLELRKVPPELNNISKLNEHFSRFGTLV NLQVAYNGDPEGALIQFATYEEAKKAISSTEA VLNNRFIKVYWHREGSTQQLQTTSPKVMQPL VQQPILPVVKQSVKERLGPVPSSTIEPAEAQS ASSDLPQVLSTLLA*QKQCIIQLL/WKAAQKT LLVSTSAVDNNEAQKKKQEALKLQQDVRKR KQEILEKHIETQKMLISKLEKNKTMKSEDKAE IMKTLEVLTKNITKLKDEVKAASPGRCLPKSI KTKTQMQKELLDTELDLYKKMQAGEEVTEL RRKYTELQLEAAKRGILSSGRGRGIHSRGRGA VHGRGRGRGRGRGRGVPGHAVVDHRPRALEIS

SEQ ID	SEQ ID	Met	CEO	Daniel I	1 6 - 3 3 3	
NO: of	NO: of	hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	1 -	/=possible nucleotide deletion, \=possible
				sequence	1	nucleotide insertion
						AFTESDREDLLPHFAQYGEIEDCQIDDSSLHA
		}	ŀ		1	VITFKTRAEAEAAAVHGARFKGQDLKLAWN
			<u> </u>		<u> </u>	KPVTNISAVETEEVEPDEEEQREIIIA
506	1856	A	3911	1952	919	DAELSGTLSLVLTQCCKRIKDTVQKLASDHK
	1	ļ		1	1	DIHSSVSRVGKAIDKNFDSDISSVGIDGCWQA
		l		İ		DSQRLLNEVMVEHFFRQGMLDVAEELCQES
				l		GLSVDPSQKEPFVELNRILEALKVRVLRPALE
		l		İ	ļ	WAVSNREMLIAQNSSLEFKLHRLYFISLLMG
					Ĭ	GTTNQREALQYAKNFQPFALNHQKDIQVLM
ļ						GSLVYLRQGIENSPYVHLLDANQWADICDIFT
1					i.	RDACALLGLSVESPLSVSFSAGCVALPALINIK
			1			AVIEQRQCTGVWNQKDELPIEV\DLG*KSAGY
ĺ		[[HSIFACPILRQQTTDNNPPMKLVCGHIISRDAL
507	1857	A	3936	439	18.	NKMFNGSKLKCPYCPMEQSPGDAKQIFF
307	1037	A	3930	439	18.	SHPFSPAPGICPDAPPPLPRPSKGLGHPGTAGA
		ĺ		ĺ	ĺ	PGSGARCHPPSTCSPSWASPG*GAKASPALPR
						SHGVTLLCKAQAHLCRGEDSKDASGSTSQA WEPG*GAWGMPRCQGPALGSCFCPPGTTVO
					,	· · · · · · · · · · · · · · · · · · ·
508	1858	A	3944	120	412	RPAKQRDKRNRHLGR WCPAGTLDFPGPQEMVLLEIEVMNQLNHRNL
	1050	'`	3544	120	712	IQLYAAIETPHEIVLFME\YECPK*W*GLGGGT
	4		· ·			TRHGASRGGVCAHSIEGGELFERIVDEDYHLT
						EV EV
509	1859	A	3949	31	392	LTKTPSPREKGRGVLSVLLMMI*KCRVIFVKIP
		ľ				MVFFLQNFC/RIILNVA\WTGD*PNTL*KEORG
						ITFSDSKS*YKATKIKTMWYCHKNRYID/ERN
						RIEIPEINPCICDKIIFRKLSMTTQ
510	1860	A	3954	1013	885	FSETRACCPRLEHSGRIEAHCSLNIPGSSDPPT
1				•		SASSVAATTG
511	1861	Α	3956	1	1054	PPAWAPRSPLIWAPTSGRHPCRAALPWSTSSV
ŀ						RWQPSEKQPPPPAHRGPADSLSTAAGAAELS
İ						AEGAGKSRGSGEQDWVNRPKTVRDTLLALH
ł	i			1		QHGHSGPFESKFKKEPALTAVARTARKRKPS
-						PEPEGEVGPPK\TTERPSRGCPHPQRGSRSP*L
				İ		LHPLLCLRHHPLPHLIPTGPHRLKRPRM\P\SP
ł						MAALILVADNAGGSHASKDANQVHSTTRRN
}				,		SNSPPSPSSMNQRRLGPREVGGQGAGNTGGL
	- 1			İ		EPVHPASLPDSSLATSAPLCCTLCHERLEDTH
				ļ		FVQCPSVPSHKFCFPCSRQSIKQQGASGEVYC PSGEKCPLVGSNVPWAFMOGEIATILAGDVK
į						VKKERDS
512	1862	A	3957	1086	3	QDRARLDCSSATSAHCNLRLPGS*DSPASASR
		•	-,,,	-500	-	VAGTTDTHHHTWLILGSSVQTGFDHVGQAG
1						LELLTSGDPPISASESAGIMGMSHCVWP*SWG
					Ī	LSHHMAPPQGDGGRARGTPGPEQSFWNLSC
						H*PRCQVPS*LMTQL/FWGRHQY:NPTMKRGK
	1		1	ļ		LRHREACSLPLPGEGEPGLQPSS*SQNPCSSPL
]	i		1		FHHGL*AWLWCPELLLQGQARRH*RSPPS/FK
	l	İ			İ	CPATLSLTAWSQTKRLRSQFLLLPWL*RAL*H
ļ	j		J	J	J	PP\CHWPSRRSLGDPLLPRSQG*RDGT*ASTFC
						SYF*DTESHLVAQAGVQWRDLGSLQPPCPRL
1	1		ļ	l		K\RFSRLSPPSSYTHRYVPSHLAESCISSRDRIP
			i i		İ	PSRPDRSRNSNSLSR
	-	- 1	i			
513	1863	A	3961	3038	476	
513	1863	A	3961	3038	476	VALTTSMCCNKQVIVIDKIKSASIADRCGALH
513	1863	A	3961	3038	476	VALTTSMCCNKQVIVIDKIKSASIADRCGALH VGDHILSIDGTSMEYCTLAEATQFLANTIDQ VKLEILPHHQTRLALKGPDHVKIQRSDRQLT
513	1863	Ā	3961	3038	476	VALTTSMCCNKQVIVIDKIKSASIADRCGALH
513	1863	A	3961	3038	476	VALTTSMCCNKQVIVIDKIKSASIADRCGALH VGDHILSIDGTSMEYCTLAEATQFLANTIDQ VKLEILPHHQTRLALKGPDHVKIQRSDRQLT

SEQ ID NO: of nucleotide peptide contide sequence (A-Alamine C-Cysteline, nucleotide location peptide contide sequence (A-Alamine C-Cysteline, 1944) DNO: mucleotide location (99496) SISTN (1946) DNO: mucleotide location (99496) SISTN (1946) SISTN	SEC IF	CEO ID	3.4.4	1000	D 1	T No. 20 1	1. (
nucleotide seq- uence verse of the sequence verse of the sequence verse of the sequence verse of peptide sequence verse of			1		1		Amino acid sequence (A=Alanine C=Cysteine,
Sequence	1		1 1100			· ·	
uence unce operation of the property of the pr		1		1			
uenice	1		l				
amino scid residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence per sequence peptide sequence per		"""				-	O-Glutomine, N-Asparagine, P-Proline,
reaidue of poptide sequence	Licitor		ļ	714	, .		C-Clutamine, K=Arginine, S=Serine,
peptide	1	}	ì	ł	t .		
				ţ		sequence	
SIASSTYGLAGGVYRITETIEVVLTADPYTOF GIQLOGSVATATILSPARENCO VLQIGDRVMAJNGIPTEDSTEEASQLLRDSSI TSKVTLEIEPDVAESVPILSVEANSPARENCO VLQIGDRVMAJNGIPTEDSTEEASQLLRDSSI TSKVTLEIEPDVAESVPISSGTEFIVKLEKKANDVAILOGOE EDLYKLRIKEDEDISDGESGAITYTELKR GIELOBKLLAIDMRIDNCSMEDAVQILOGOE EDLYKLRIKKEDEDISDGESGAITYTELKR YGGPLGUTISGTEEPPDL*IDSLTKGGLAERT GAIHLIGDRILAINSSS,KGPLSEAHLLQMAG ETVTLKIKKQTDAGSASSPKKFPISSHLDQLAGO VEEDSSPAQKPGKLSSMYPSHGCPSVDSAVD WSSPAGSAUDTSYGTEGTSFQASGYNFNTYD WSSPAGSAUDTSYGTEGTSFQASGYNFNTYD WSSPAGSOLPSFGRESSSRPHYSQATE WSSTASGFAGAADSASTRCEENFYSQATE WSSTASGFAGAADSASTRCEENFYSQTTESNTLPS DVGRESVTLRKMKQGEEMSFTPVSQLSSED WDSSTASGFAGAADSASTRCEENFYSQTTESNTLPS DVGRESVTLRKMKQGEEMSFTPVSQLSSED DVGRESVTLRKMKQGEEMSFTPVSUHKVT LYKDSDMEDFGFSVADGLLEKGYYVKNIRPA GFGDLGGKRYDRLQVHVRNTEDFDCLV VPLIAESGNKLDLVISRPILASQKSIDQOSLPG PSEQUSAFFQOPSHGUEEMSFTPVSUHKVT LYKDSDMEDFGFSVADGLLEKGYYVKNIRPA GFGDLGGKRYDRLQVHVRNTEDFDCLV VPLIAESGNKLDLVISRPILASQKSIDQOSLPG DSEQUSAFFQOPSHGUEEMSFTPVSUHKVT LYKDSDMEDFGFSVADGLERGNSTTPVSUHKVT LYKDSDMEDFGFSVADGLERGNSTTPVSUHKVT LYKDSDMEDFGFSVADGLERGNSTTPVSUHKVT LYKDSDMEDFGFSVADGLERGNSTTPVSUHKVT LYKDSDMEDFGFSVADGLERGNSTTPVSUHKVT LYKDSDMEDFSVADGLERGNSTTPVSUHKVT LYKDSDMEDFSVADGLERGNSTTPVSUHKVT LISTOCH	1	1	1	ľ		ļ	
GIQLOGSYATETLSSPPLISYIEADSPAERCE VLQIDDEWMAINGIPTEDSTEEASQLERDSSI TSKVTLEIEFDVAESVIPSSGFFIVKLPKKHN VELGITISSPSSRKPOPLVSIDIKKGSVARRT GTLEIGDKLLAIDNIELDNCSMEDAVQILQOC GELVVKLIKRDEDBNSDQESSGAIPYVELKR YGGPLGUTISGTEFPPID-19ISLTKGGLAERT GAHHGBRILAINSSK KGKPLSSAHLLQMAG ETVTLKIKKOTDAQSASSPKKFPISSHLBJLGD VEEDSSPAQKPGKLSDMYPSHGCPSVDSAVD SWDOSAIDTSYGTEGTISRQASQTWFNTYD WRSPKQRGSUSPYTRYBYGDYDLOSYED WRSTASGRAGAAUDSATEQEBETWSQALE DLETCQQSGLRELEATIMSGSTMSLNBEAPT PRSPAGSORPSFGERSSSRPHYSQTTRSNTLPS DVGRKSVTLRKMKQEIKEIMSPTPVELHKVT LYKDSDMEDPGFSVADGLILEKGVYYKINRPA GRÖDLGGLKYYRRLQVYHVRTRDPDCCLV VPLLAESGNKLDVISRPLASQKSIDQQSLPG DPSEQNSAFFQQPSHGGNLETREPTIVIL LEKQQVSGMATKRLARQCILIBREXISHANG NLGRSKSKQLFDYLVIDFESTCWNDGKHHH SQEITEPAVLLNTSTGDISSPGAYCYQPGHPI LSFCCMELTGIKQAQVDEGVYLKICLSQFCK WHKIQQQKNIB-ATGIGDSEPGAYCYQPGHPI LSFCCMELTGIKQAQVDEGVYLKICLSQFCK WHKIQQQKNIB-ATGIGSFSPGAT-SSKNCICYL VR-RISYTY-SKHKSKGC CFFWGISTHCDTCDFSPQTTEG**EGDLWSL DLCPEPELARKPLFKTKTYQSTF*SISKNCEPTC NFILEEGTOLLIPVQVKHNPCHRLTPEGOTVQL NADS 182 182 CFFWGISTHCDTCDFSPQTTEG**EGDLWSL DLCPEPELARKPLFKTKTYQSTF*SISKNCEPTC NFILEEGTOLLIPVQVKHNPCHRLTPEGOTVQL NADS SMCCALSWYTHERAKADDXSMFVRIKTLGT GAFGEVCLARKVDTKALYATKTLRKDVLL RNQVAHVKAERDILAEADNEWVVRLYYSFQ DEDNLYVMDYPFGGDMSSLLKRMGFPESL ARFYLAELTCAVESVHKMGFHERDIKFDNLLD ROGHIKLTDFGLCTGRYWTHDSKYYQSGDHP RQDSMDFSNEWGDPSSCRCGDRLKPLERTAA RQNGMDFSNEWGDPSSCRCGDRLKPLERTAA RQNGMDFSNEWGDPSSCRCGDRLKPLERTAA RQNGMDFSNEWGDPSSCRCGDRLKPLERTAA RQNGMDFSNEWGDPSSCRCGDRLKPLERTAY RCHPHILTPTICATYQC CDWWSVGVILFEMLVGQPFFLAGTYCSGDPD DRACKNGADEIKAHPIFNOFDPSQ*FEDSS AFKQFPNITTPTICATYTQTQL CDWWSVGVILFEMLVGQPFTLAGTYTQTQL CDWWSVGVILFEMLVGQPFTLAGTYTCCV VFLDRVPLCHPGWSAVVQSCDPD DNTCSLIBNTRDLLVV VFLDRVPLCHPGWSAVVQSCDPD DNTCSLIBNTRDLVVV VFLDRVPLCHPGWSAVVQSCDPD DNTCSLIBNTRDLVVV VFLDRVPLCHPGWSAVVQSCDPD DNTCSLIBNTRDLVVP VFLDRVPLCHPGWSAVVQSCDPD DNTCSLIBNTRDLVP SCRCEPTINTOTAGPVCHARTSTVLRVIS AFKCPPHILTPTTCAGVCACHFTH VQAGLELLTSGOLPALASQSAGITGSISRAR PENOFERINF SCRS.PFLOYNHYPCLANSVFSVEMGFHH VGQAGLELLTSGOLPALASQSAGITGSISRAR STEPPLOYNHYPCLANSVFSVEMGFHH VGQAGLELLTSGOLPALASQSAGITGSI		 	├	ļ	sequence		
V.Q. GDRVMAINGIPTEDSTEEASQLLRDSS TSKVTLEIEPVAESVPSSTFHVLKPKHN VELGITISSPSSTRYGOPL VISDIKKGSVAHRT GTLELGBLALDNILDNOSMEDAVOILQOC EDLVKLKIRKDEDNSDQESSGAHYTVELKR YGGPLQVITISGTEPPUP-U VISSLTRGGLAERT GAHHGDRILAINSSSLKGRPLSSAHLLQNAG ETVTLKIKKOTDAGSASSPKKFPISSHLSDLGD VEEDSSPAQRFGKLSDMYPSHGCPSVDSAVD SWDGSALDTSYGTEGTSSPAGSGVNFNTYD WRSTASGFAGAADSAGTVGENFSTYDGAVD SWDGSALDTSYGTEGTSSPAGSGVNFNTYD WRSTASGFAGAADSAGTGEENFWSQALE DETCOGGGILRELATIMSGSTIMSLHREAPT PRSPAGSDRPSFQERSSSRPHYSQTTESNTLPS DVGRRSVTLRKMKGCREMBSPTPVELKKVT LYKDSDMEDFGFSVADGLLERGVYVKNIRRA GFÖDLGGKRYPRILQVNHVRTRDFDCLV VPLIAESGNKLDLVISRPLASOKSIDQQSLPG GFÖDLGGKRYPRLUVHFSTCWDOKHHH SQEIEPPAVLLNTSTGQIDEEFQAVVQPQEHPI LSFECMELTGIKQAQVDEGVPLKICLSQPCK WHKLQQQNIPFATGISEPSOFFSKENCCYL VRNSTYTYSKIKSKOC URNSTYTSKIKSKOC CRFWGISTHCOTCDPLSSPQTTEG**EGDLWSL DLLCPEFLARRFFKTKTYQSTF*SISKNEFTC DLLCPEFLARRFFKTKTYGSTF*SISKNEFTC STATE STAT	1		1	ŀ			SLASSIVGLAGQVVHIEITEVVLTADPVTGF
TSKYTLEIEPDVAESVIPSSGTFHVKLPKKHN VELGITSSPSSRKPDPLVISDIKKSSVART GTLELGDKLLAIDNIRLDNCSMEDAVQILQDC EDLVVKLIKKDEDBNSDEQSSGAIPTVPLKR YGGPLGVTISGTEPPPDL*IDSSLTKGGLAERT GAIHIGDRILAINSSLKGPLSSAIHLLQNAG ETVTLKIKKQIDAQSASSPKCPISSLADLIGD VEEDSSPAQRPGKLSDMYPSHGCPSVDSAVD SVBOSSADIDTSYGTEGTSPQASGVYNFNYD WRSPKQRGSUSPYTKPRSQTYPDVGLSYED WRSPKQRGSUSPYTKPRSQTYPDVGLSYED WRSPKSGRFAGSALDSAFTGEERFWSQALE DLETCGQSGIRRELEATIMGSTMSLINEAPT PSFAGSDRPFFGERSSSRPHYSQTTRSNTLPS DVGRKSVTLRKMKQEIKEIMSTPVELHKVT LYKDSDMEDPGFSVSSSRPHYSQTTRSNTLPS DVGRKSVTLRKMKQEIKEIMSFTPVELHKVT LYKDSDMEDPGFSVSSSRPHYSQTTRSNTLPS DVGRKSVTLRKMKQEIKEIMSFTPVELHKVT LYKDSDMEDPGFSVSSSRPHYSQTTRSNTLPS DVGRKSVTLRKMKQEIKEIMSFTPVELHKVT LYKDSDMEDPGFSVSSRSPHYSQTTRSNTLPS DVGRKSVTLRKMKQEIKEIMSFTPVELHKVT LYKDSDMEDPGFSVADGLIERKSTMSATANG GFGDLGGLRFVDRLJQVHVMTRDPDCLV VFLIASSGNKLIDVISMPIASKSIDQSLPG PSEQNSAFFQQPSHGGNLETREPTNTL SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLLINTSTOGIDSEPGAYVQPQEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTALSVEHPI SEGELEFFAVLTATSVEHPI SEGELEFFAVLTATSVEHPI SEGELEFFAVLTATSVEHPI SEGELEFFAVLTATS	1	[ł	l	l	ł	GIQLQGSVFATETLSSPPLISYIEADSPAERCG
VELGITISSPSSKRODPL VISDIKKGSVAHRA GTLELIGDKLLAIDINELDNCSMEDAYQILQQC EDLVKLKIRKDEDNSDEQESSGAITYTVELKA GELVKLKIRKDEDNSDEQESSGAITYTVELKA YGGPLQVITISGFEFFFFF10.1*NSLTKGGLAERT GAHIGDRILLAINSSSLKGKPLSEAIHLLQMAG ETVTILKIKQDAQASASPKKFPISSHDLGD VEEDSSPAQRPGKLSDMYPSHGCPSVDSAVD SWDOSAIDTSYGTEGTSFQASGYNFNTYJO* WRSPKQRSULSYTVKRPSGTYTDVGLSYED WDRSTASGFAGAAUSABTEQEENFWQALE DLETCOQSGILRELEATINGSGTASISLHEAPT PRSPAGSDRPSPQERSSSPHYSGTTRSNTLG DVGRKSVTLRKMCQEKEMBPTPVELHKVT LYKDSDMEDFGFSVADGILEKGVYVKNIRPA GPODLGGLKPYDRLLQVNINVRTDPCCLV VPLIAESGNKLDLVISRNFLASQKSIDQQSLPG DYSEGNSAFFQQPSIGGNLETREPDTCLV VPLIAESGNKLDLVISRNFLASQKSIDQQSLPG DYSEGNSAFFQQPSIGGNLETREPDTCLV VPLIAESGNKLDLVISRNFLASQKSIDQQSLPG DYSEGNSAFFQQPSIGGNLETREPDTCLV VPLIAESGNKLDLVISRNFLASQKSIDQQSLPG DYSEGNSAFFQQPSIGGNLETREPDTVLIT SGEIEFFAVLLNTSTQGIDERRYSIAPANG NLGRKSKQLPUJVIDFFSTXDAGHH SGEIEFFAVLLNTSTGQIDERRYSIAPANG NLGRKSKQLPUJVIDFFSTXDAGHH SGEIEFFAVLLNTSTGQIDERRYSIAPANG NLGRKSKQLPUJVIDFFSTXDAGHH SGEIEFFAVLLNTSTGQIDERRYSIAPANG NLGRKSKQLPUJVIDFFSTXDAGHH SGEIEFFAVLLNTSTGGIDERRYSIAPANG NLGRKSKQLPUJVIDFFSTXDAGHH SGEIEFFAVLLNTSTGGIDERPQTALYCICLSQPFQ WHKLQQQKNIHATGISENDFYSIKMCCVL VRYRISTTY*SKHESKOC STEMELSTANLANTSTGGIDERPQTALYCICLSQPFQTALYCICLSQPFQTALYCICLSQPFQTALYCICLSQPFQTALYCICLSQPFQTALYCICLQPTPLTTCATAGATAGATAGATAGATAGATAGATAGATAGATAGA			ì	j			VLQIGDRVMAINGIPTEDSTFEEASQLLRDSSI
GTLELGDKILAIDNRILDNESMEDAVQILOGE EDLVKIKIRKDEDNISDEGESSGAINTYVELKR YGGPLGITTIGTEEPFDI	1						
EDLVKLKIRKOEDNSDEGESSGAITYTUELK GAPICAUTISGEEPFPOIL*INSTAKGGLAERT GAHIGIDRIL\(\)\(\)\(\)\(\) GAPT GAHIGIDRIL\(\)\(\)\(\)\(\)\(\)\(\)\(\)\(\)	j	ł	ļ	ļ	ļ	j	
		1	1]	1		
GAHHGDRILVAINSSSLKGKPLSEAHHLLQMAG FLYTLKIKKQTDAGASSPKKPRISBLDLGD VEEDSSPAQKPGKLSDMYPSHGCPSVD8AKI) VEEDSSPAQKPGKLSDMYPSHGCPSVD8AKI) VEEDSSPAQKPGKLSDMYPSHGCPSVD8AKI) VEEDSSPAQKPGKLSDMYPSHGCPSVD8AKI) VEEDSSPAQKPGKLSDMYPSHGCPSVD8AKI) VEEDSSPAQKPGKLSDMYPSHGCPSVD8AKI VEEDSSPAGSURSPTKPSCTYPDVGLSYED WDRSTASGF AGAAUSAETEQEENFWSQALE DLETCOQSGILKELBATIMSGSTNSINHEAPT PRSPAGSDRPSPGERSSSPHYSCTTRSNTLS DVGRKSVTLRKMKQEIKEMBPYPELHKVT LYKDSDMEDFGFSVADGILEKGVYVKNIRPA GRODLGGIKPYDRLQVNIVRTPDPCLV VPLIAESGNKLDLVSINPHAAGKSIDQQSLPG D*SEGNSAFPQQSSIGGMLETREPTNTL VPLIAESGNKLDLVSINPHAAGKSIDQOSLPG VPLIAESGNKLDLVSINPHAAGKSIDQOSLPG VPLIAESGNKLDLVSINPHAAGKSIDQOSLPG VPLIAESGNKLDLVSINPHAAGKSIDQOSLPG VPLIAESGNKLDLVSINPHAAGKSIDQOSLPG VPLIAESGNKLDLVSINPHAAGKSIDQOSLPG VPLIAESGNKLDLVSINPHAAGKSIDQOSLPG VPLIAESGNKLDLVSINPHAAGKSIDQOSLPG VPLIAESGNKLDLVSINPHAAGKSIDQOSLPG VPLIAESGNKLDLVSINPHAAGKSIDQOSLPG VPLIAESGNKLDLVSINPHAAGKSIDAON VPLIAESGNKLDCVLVV VPLIAESGNKLDCVLVV VPLIAESGNKLDCVLVV VPLIAESGNKLDCVLVV VPRISTTY*SKHKSKGC VPLIAESGNKLDCVLVV VPRISTTY*SKHKSKGC VPLIAESGNKLDCVLVV VPRISTTY*SKHKSKGC VPLIAESGNACH VPLIAE	}	ł	ł		ļ	ł	
ETYTLKIKQIDAQSASSYKKPRISSHLDLGD VEEDSSPAQKPKLSDMYPSHGCPSVDSAVD SWDGSAUDTSIYGTEGTISFQASGYNIPSHTYD WRSPKQRGSLSPYTKPRSQTYPVQLSYED WDRSTASGFAGAAUDSAFTEQENFWSQALE DLETCQQSGLRELEATIMSGSTMSLNHEAPT PRSPAGSDRSPGCRSSSPHYSQTTRSNTLPS PRSPAGSDRSPGCRSSSPHYSQTTRSNTLPS PRSPAGSDRSPGCRSSSPHYSQTTRSNTLPS DVGRKSVTLRKMKQEIKEIMSPTPVELHKVT LYXDSMEDPEGFSVADGLLERGVYVKNIRPA GPGDLGGLKPYDRLLQVSHVKNTRDFDCCLV VPLIAESONKLDLVISRIPLASQKSDQQSLPG DPSEQNSAFFQQPSHGGNLETREPTNTL 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSJAPANG MLGRSKSQLDVYLVIDPISTOMDGKHHH SQEIIEFPAVLLNTSTGQIDSEFQAYVQPQSHPI LESFCMELTIGKQAQVDEGPYLKICLSQFCK WIHKIQQQKNIBFATGISEPSDF*SKIMCICYL VR*RISYTY*SKIKSKGC 182 CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL DLLGPEFLARRPLFKKTYQSTP*SISKNEFTC PNPIEEGTDLIP*QVKHNPCHRITPEEGTVQL NRADS 1866 A 3977 2 1337 KMLC/OKESNYIRLKRAKMDKSMFVKIKTLGI GAFGEVCLARKVDTKALVATKTLRKKDVLL RNQVAHVKAERDILAEADNEWVALLYYSFQ DKDNLYFVMDYIPGGDMMSLLRMGIFPESL ARFYLAELTCAVESYHKMGFHRDIKFDNLID RDGHIKLTDFGLCTGFRWHDSKYYQSGDHP RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLYGTNVIAPGSKDV CDWSVGVILFEMLVQQPFFLAOTPLETOM KVINWOTSLHPPQAKLSPSABDLIIKLCRGPPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFF*NHTTPTDTISNFDPVUPPIKLWSDDN EEENVADTLNGWYKNGKHEHAFYETFRRF FDDNGYPYNYPKPIEVSYNSQGSEQQSDEDD QNTGSEIKNDLVY 517 1867 A 3980 1358 1022 FFFKKFTQSLGFLESSFLFSCFFFFFIREVLFCY VCLORVSVGVILFEMLVQOPPFLAOTPLETOM KVINWOTSLHIPPQAKLSPSABDLIIKLCRGPPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFF*NHTTPTDTISNFDPVUPPIKLWSDDN EEENVADTLNGWYKNGKHEHEAFYETFRRF FDDNGYPYNYPKPIEVSYNSQGSEQQSDEDD QNTGSEIKNDLVV FFFKKFTQSLGFLESSFLFSCFFFFFIREVLFCY VCLORVSVGVILFEMLVQOPPFLAOTPLETOM KVLORVSPCKLARGVANDSSSEQGSDEDD QNTGSEIKNDLVV FFFKKFTQSLGFLESSFLFSCFFFFFIREVLFCY VCLORVSVGVILFEMLVQOPPFLAOTPLETOM KVLORVSPCKLARGVANDSSSEQGSDEDD QNTGSEIKNDLAVQQPPRVLOLQV FFFKKFTQSLGFLESSFLFSCFFFFFIREVLFCY VCLORVSVGVILFAGNAVQNGVASTELGKSHHHS 518 1869 A 3986 974 666 SPEMESHETIQAGVQWHHLSSLQPLPPGFK*F SCFSLFPSLGYRRVSPCLANSVSVEMGGHLH VGQGGLELLTSGQLPALAGQAGTGSHFAR PENGEENF 519 1869 A 3994 751 126 NGGRTAUSWEDGHCLLLLPLGTGCLGKFIVV		ŀ					YGGPLG\ITISGTEEP\FDL*IISSLTKGGLAERT
VEEDSSPAQKPGKLSDMYPSHGCPSVDSAVD SWDGSAIDTSVGTGGTSGGVTNNTNTVD WRSPKQRGSLSPVTNKPRSQTVPDVGLSYED WDRSTASGFAGAA\DSAETQEENIFWSQALE DLETCQGSGLRELEATIMSGSTMSLNHEAPT PRSPAGSDRPSFQERSSRPHYSQTTRSNLPS DVGRKSVTLRKMKQEKEKIMSPTPVELHKVT LYDDSDMEDPGFSVADGLLEKGVYVKNIRPA GPGDLGGLKPYPDRLLQVHNVRTRDFDCCLV VPLIAESGNKLDLVISRNPLASQKSIDQQSLPG D*SEQNSAFPQQPSHGGNLETREFITNIT SQEIEFPAVLLNTSTGGDISEFQAYVQPQEHPI LSEFCMELTGIKQAQVDEGVPLKICLSQFCK WIHKIQQQKNIB*ATGISEFSQAYVQPQEHPI LSEFCMELTGIKQAQVDEGVPLKICLSQFCK WIHKIQQQKNIB*ATGISEFSQAYVQPQEHPI LSEFCMELTGIKQAQVDEGVPLKICLSQFCK WIHKIQQQKNIB*ATGISEFSQFY*SKIMCICYL VR*RIS*TY*SKIKKSKQC 515 1865 A 3969 492 182 CRFWGIS*HCDTCDPJSPQTTEG**EGDLWSL DLLGPEFLARKPLFKTKTYGST*SISKNEFTC PNFILEEGTDLIF*QVKHNFCHRLTPEEGTVQL NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYWRLKRAKMDKSMFVKIKTLQT GAFGEVCLARKVDTKALVATKTLRKKDVLL RNQVAHVKAERDILAEADNEWVYLLYYSFQ DKDNLYFVMDYPGGDDMSLLIRMGIPPESL ARFYIAELTCAKESVLKMFGHRDIKFDNILLD RDGHIKLTDFGLCTGFRW*HDSKYYQSGDHP RQDSMDFSNEWGDPSSCCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGYILFEMLVQQPPFLAQTIFLETQM KVNWGTSLHPPQAKLSPRASDLIIKLCRGPP DRLGKNGADEIKAHPI*POQRDSPSCCGDRLKPLETRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGYILFEMLVQQPPFLAQTIFLETQM KVNWGTSLHPPQAKLSPRASDLIIKLCRGPP DRLGKNGADEIKAHPI*POQRDSPSCCGDRLKPLETQMF KVNWGTSLHPPQAKLSPRASDLIIKLCRGPP DRLGKNGADEIKAHPI*POQRDSPSCPGDSSS AFKQPP*NHTTDTISNEPDVDPDKLWSDDM EEENWADTLUNGWYKNGKHPEHAFYEFTFRF FDDDNGYPYNYPKPRIEYEYNSQGSEQQSDEDD ONTGSBIKNRDLVYV FFFKKFTQSLGFLLFSSFLFSCFFFFHFVLFCY VFLDRYPLCHPGWSAVVQSQVTVNLPPSWD *RCRPPHLANLCNFCRDSFTTLPRLVLNTWA QAFGPPPRYLLQLQV **SPEMSSHTTQAGVQWHHLSSLQPLPPGFK*F* SCFSLPP-LGYRHVPPCLANSVFSVEMGWGHH VGQAGLELLTSGDLPALASQSAGTTGSHRAR PENGFENF 519 1869 A 3994 751 126 1126 1127 1126 1128 1129 1126 113NQGHNALWKAAGPPLKAGYCQSFSPC DSLKVJOSWDEKDLTVYQRDTHKRSXLXWIS QRGKLAVEMEEGHCLLLLPLJGTCLGIKCPV						1	GAIHIGDRIL\AINSSSLKGKPLSEAIHLLQMAG
SWDOSAIDTSYGTEGTSFQASGYNINTYD WRSPKQRGSLSPYNTRRSTYDVGLSYED WRSTAGGFAGAADSAFTEQEENFWSQALE DLETCQSGILRELEATIMSGTMSLNHEAPT PRSPAGSDR*SFGPERSSSR*PHYSQTTRSNTLPS DVGRKSYTLRKMKQEIKEIMSPIPVELHKVT LYKDSDMEDFGFSVAGGILEEKGYVYKNIRPA GPÖDLGGLKPYDRLLQVHYWTRTRDFDCLV VPLIAESORKLDLVISRNPTASQKSIDQQSLPG D*SEQNSAFFQQPSHGGILETREPTIVTI 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSIAPANG MLGRSKSKQLFDYLLVIDFSTCWNDGKHHH SQEIIEFPAVLINTSTGQIDSEFQAYVQPGHPI LSEFCMELTGIRQQAVDEGPYLKICLSQFCK WIHKIQQQKNIIFATGISEPSDF*SKIMCICYL VR*RISYTY*SKIKSKGC 1515 1865 A 3969 492 182 CRF WGISTHCDTCDFLSPQTTEG**SGDLWSL DLLGPEFLARRPLFKTKTYQSTF*SISKNE/FTC PNFIEEGTDLIP*QVKHNCHRITPEEGTVQL NRADS 166 A 3977 2 1337 KMC/JCKESNYIRLKRAKMDKSMFVKIKTLGI GAFGEVCLARKVDTKALYATKTLRKKDVLL RNQVHYKAREDILAEADNSKLYVGLYSFG DKDMLYFVMDYPPGGDMMSLLIRMGIPPESL ARFYIAELTCAVESYHKMGFHRDILEPDMLID RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDPSSCCCGDRLKPLERRAA RQHGCLAHSLVGTHVNLAPGHASKYQSGDHP RQDSMDFSNEWGDPSSCCCGDRLKPLERRAA RQHGCLARSLVGTHVNLAPGHASTCHLTGYTQL CDWWSVGYILFEMLVQOPPFLAGTPLETQM KVINWGTSLHIPPQAKLSPRSJUKLLCGFGPE DRIGKNGADEIKAHPI*PNQFDFSQ*PEDSSS AFKQFP*NHTTPTDTSNFDPVVDPDKLWSDDN EEENVADTLNGWYKNOKEHPLAFVETTFRFF FDDNGYPYNYPKPLEVEYNSQGSEQQSDEDD QNTGSLIKNADLVY 517 1867 A 3980 1358 1022 FFFKKFTOSLGFLESTSFLFSCFFFFHFVLFCY VFLDRYPLCHPGWSAVVQSQTVANLPPSWD *RCRPPHLANLCNCCROSFTTLPRLVLNTWA QAIPQPQPPKVLCLQV *FFKKFTOSLGFLLSTSFLFSCFFFFFHFVLFCY VFLDRYPLCHPGWSAVVQSQTVANLPPSWD *RCRPPHLANLCNCCROSFTTLPRLVLNTWA QAIPQPQPPKVLCLQV 518 1869 A 3994 751 126 NQGRRYGLGRTLASQGAGITGSHRAR PENGFENIF 519 1869 A 3994 751 1126 NGGRRYGLGRTCLVNQMPASSILGKSHHHS LISNQGHNALWKAAGPLPLKAGYCQSFSPC DSLKVJOSWDEKDLTVYQRDTHKRSVLKWBS QRKKLAVEMBEGGHCLLLLPLLJCTGCLGIKCPV			J			j	ETVTLKIKKQTDAQSASSPKKFPISSHLSDLGD
WRSPKORGSUSPYTIKPRSGTYPDVGLSYED WDRSTASGFAGAADSASTEGENFWSQALE DLETCGQSGLRELEATIMSGSTMSLNHEAPT PRSPAGSDRPSFQERSSSRPHYSGTTRSNLHEAPT PRSPAGSDRPSFQERSSSRPHYSGTTRSNLPS DVGRKSYTLRKMKQEIKEIMSPTPVELHKYT LYKDSDMEDFGFSVADGLLEKGYVYKINRPA GPGDLGGLKPYPDRLLQVHNIRTRDFDCCLV VPLIAESGNKLDLVISRPPLASQKSDQQSLPG D*SEQNSAFPGQPSHGGNLETETPITTI 514 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSIAPANG NLGRSKSQLEPYLLVIDFESTCWNDGKHHH SQEIEIFPAVLINTSTQGIDSSFQAYVQPQEHPI LSEFCMELTGIKQAQVDEGVPLKICLSQFCK WINKIQQQKINEATGISEPSFCMMCICYL VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL DLLGPEFLARRPLFKTKTYQSTF*SISKNEFTC PNPILEEGTDLIP*QVKHNPCHRLTPEEGTVQL NRADS 516 1866 A 3977 2 1357 KMLCVQKESNYIRLKRAKMDKSMFVKIKTLGI GAFGEVCLARKVDTKALYATKTLRKDVLL RNQVAHVKAERDILAEADNEWVVRLYYSFQ DKDNLYFVMDYPEGGDMMSLLIRMGIPFESL ARRYIAELTCAVESVHKMGHINEKPDNILLD RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSSMPSNEWGDPSSCRCGDRLKPLERDAA RQHGRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGVILFEMULQOPPFLAQTPLETQM KVNNWGTSLHIPPQAKLSPEASDLIIKLGRGPP DRLGKNGADEIKAHPI*PNQPDSCYSGDHP PDRGYPYMYPKPLEYSYNSQGSQGDEDD NTGSEIKNRDLYVV VFLDRYPLCHFEWSAVVQSQTTVNLPPSWD **RCFPPHAANLCNFCRDSFTTLPRLVLNTWA QATGPPQPPRV LGLQV VFLDRYPLCHFWGSAVVQSQTTVNLPPSWD **RCRPPHLANLCNFCRDSFTTLPRLVLNTWA QAGGLELLTSGDLPALASQGGTIGSHRAR PENGFENF 519 1869 A 3984 751 126 NGGRRFGLGRTLVNQMPASSLGKSHHS LISINQGHNALWKAAGPLPLKAGYCQSFSPC DSLKYGNSWEKDLTYPQRDTHKRSVLWMS QRKKLAVEMBEGGHCLLLLPLJGTGCLGIKKPIV							VEEDSSPAQKPGKLSDMYPSHGCPSVDSAVD
WRSPKORGSUSPYTIKPRSGTYPDVGLSYED WDRSTASGFAGAADSASTEGENFWSQALE DLETCGQSGLRELEATIMSGSTMSLNHEAPT PRSPAGSDRPSFQERSSSRPHYSGTTRSNLHEAPT PRSPAGSDRPSFQERSSSRPHYSGTTRSNLPS DVGRKSYTLRKMKQEIKEIMSPTPVELHKYT LYKDSDMEDFGFSVADGLLEKGYVYKINRPA GPGDLGGLKPYPDRLLQVHNIRTRDFDCCLV VPLIAESGNKLDLVISRPPLASQKSDQQSLPG D*SEQNSAFPGQPSHGGNLETETPITTI 514 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGLIRRKSIAPANG NLGRSKSQLEPYLLVIDFESTCWNDGKHHH SQEIEIFPAVLINTSTQGIDSSFQAYVQPQEHPI LSEFCMELTGIKQAQVDEGVPLKICLSQFCK WINKIQQQKINEATGISEPSFCMMCICYL VR*RISYTY*SKHKSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL DLLGPEFLARRPLFKTKTYQSTF*SISKNEFTC PNPILEEGTDLIP*QVKHNPCHRLTPEEGTVQL NRADS 516 1866 A 3977 2 1357 KMLCVQKESNYIRLKRAKMDKSMFVKIKTLGI GAFGEVCLARKVDTKALYATKTLRKDVLL RNQVAHVKAERDILAEADNEWVVRLYYSFQ DKDNLYFVMDYPEGGDMMSLLIRMGIPFESL ARRYIAELTCAVESVHKMGHINEKPDNILLD RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSSMPSNEWGDPSSCRCGDRLKPLERDAA RQHGRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGVILFEMULQOPPFLAQTPLETQM KVNNWGTSLHIPPQAKLSPEASDLIIKLGRGPP DRLGKNGADEIKAHPI*PNQPDSCYSGDHP PDRGYPYMYPKPLEYSYNSQGSQGDEDD NTGSEIKNRDLYVV VFLDRYPLCHFEWSAVVQSQTTVNLPPSWD **RCFPPHAANLCNFCRDSFTTLPRLVLNTWA QATGPPQPPRV LGLQV VFLDRYPLCHFWGSAVVQSQTTVNLPPSWD **RCRPPHLANLCNFCRDSFTTLPRLVLNTWA QAGGLELLTSGDLPALASQGGTIGSHRAR PENGFENF 519 1869 A 3984 751 126 NGGRRFGLGRTLVNQMPASSLGKSHHS LISINQGHNALWKAAGPLPLKAGYCQSFSPC DSLKYGNSWEKDLTYPQRDTHKRSVLWMS QRKKLAVEMBEGGHCLLLLPLJGTGCLGIKKPIV			1		-	[SWDGSA\IDTS\YGTEGT\SFQASGY\NFNTYD
DLETCGQSGILRELEATIMSGSTMSINHEAPT PRSPAGSDRPSOFRESSSRPHYSQTTRSNTLPS DVGRKSVTLRKMKQEIKEIMSPTPVELHKVT LYKDSDMEDFGFSVADGLLEKGYVYVKNIRPA GPGDLGGLKFYDRLLQVHKNTRTDFDCCLV VPLLABSGNKLDLVISRNPLASQKSIDQGSLPG D*SEQNSAFFQQPSHGGNLETREPTNTL S14 1864 A 3967 833 800 LEKQGVSGMATKRLARGLGLIRRKSIAPANG NLGRSKSKQLFDVLIVIDFESTCWNDGKHHH SQEIEEPPAVLINTSTGQDSEPCAYVQPQEHPI LSSECMELTGIKQAQVDEGVPLKICLSQFCK WHKIQQQKNIIFATGISEPS/DF*SKEMCICYL VR*RISYTY*SKHSKGC 515 1865 A 3969 492 182 CRFWGISTHCDTCDPLSFQTTEG**EGDLWSL LGPEFLARKPLFKTKTYQSTF*SISKNEFTC PNFHIEGGTDLIF*QVKHNPCHRLTPEEGTVQL NRADS S16 1866 A 3977 2 1337 KMLC/QKENSYIRLKRAKMDKSMFVKIKTLGI GAFGEVCLARKVDTKALYATKTLRKDVLL RNQVAHVKAERDILAADMSEWVVRLYYSFQ DKDNLYFVMDYIPGGDMMSLLIRMGIFFESL ARFYIAELTCAVESVHKMGFHRDIKPDNILDID RDGHIKLTDFGLTCTGFRWTHDSKYYQSGDHD RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHORCLAHSLVGTPNYDAFVYSPGDVBSC ARFVIAELTCAVESVHKMGFHRDIKPDNILDID RDGHIKLTDFGLTCTGFRWTHDSKYYQSGDLDD KVINWQTSLHIPPQAKLSPEASDLIKLCRGPE DRILGKNGADEIKAHPNGFDFSCPFEBSRS AFKQFP*NHTTPTDTSNFDPWDPDKLLRTGTYQL CDWWSVGVILFEMLVGQPFFLAQTFLETQM KVINWQTSLHIPPQAKLSPEASDLIKLCRGPE DRILGKNGADEIKAHPNGFDFSCPFEBSRS AFKQFP*NHTTPTDTSNFDPWDPDKLLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRR FDDNGYPYNYPKPIESFYTNSQGSEQQSDEDD QNTGSIENRDLVYV **RCRPPHLANLCNFCRDSTTLPRLVLNTWA QAGAGELKHPFFYNGAGSEQGSEQSDEDD QNTGSEIKNRDLVYV **RCRPPHLANLCNFCRDSTTLPRLVLNTWA QAGGELLTSGDLPALASQSAGTTGISHRAR PENGFENIF 519 1869 A 3994 751 126 NGGRHVGLCRTCLVNQMFASSILGKSHHBI LISINQGHNALWKAAGPLPLKAGYCQSFSPC DSLKYGSWDEKDLTVPQRTSVLRWIS QRGKLLAVEMEEGHCLLUPLGTECLGIKVPV]]] .			WRSPKQRGS\LSPVT\KPRSQTYPDVGLSYED
PRSPAGSDRPSPORESSRPHYSOTITENTILPS DVGRKSVTLRKMKQEIKEIMSPTPVELHKVT LYKDSDMEDFOFSVADGLEKGVYVKNIRPA GPGDLGGLKPYDRLLQVNHVRTREPDFCCLV VPLIAESGMKLDLVISRNPLASQKSIDQQSLPG D*SEQNSAFPQDPSHGDLERGFTNTL 1864 A 3967 833 800 LEKQGVSGMATKRLARQLGIRRKSIAPANG NLGRSKSKQLFDVLIVIDFESTCWNDGKHHH SQEIIEFPA-VLLNTSTGQIDSEPGAVYQPQEHPI LSEFCMELTGIKQQQVDEGVPLKICLSQPCK WHKIQQQKNIIFATGISEPS/DF*SKIMCICYL VR*RISYTY*SKHKSKGC 1865 A 3969 492 182 CRFWGISTHCDTCDELSPQTTEG**EGDLWSL DLLGPEFLARKPLFKTKTYQSTF*SISKNEFTC PNFIIEGGTDLIF*QVKHNPCHRLTPEEGTVQL NRADS 516 1866 A 3977 2 1357 KMLCQKESNYIRLKRAKMDKSMFVKIKTLGI GAFGEVCLARKVDTKALYATKTLRKKDVLL RNQVAHVKAERDILAADNEWVVRLYYSFQ DKDNLYFWDVTPGGDMUSLIRMGIPFESL ARFYLAELTCAVESVHKMGFHRDIKPDNILID RDGHIKLTDFGLCTGFRWTHDSKYVYGSGDHP RQDSMDESNEWGPPSSCRGDBLKPLERRAA RQHQRCLAHSLVGTFRYHDSKYVYGSGDHP RQDSMDESNEWGPPSSCRGDBLKPLERRAA RQHQRCLAHSLVGTPRYIAFEVLLTGTYTQL CDWWSVGVILFEMLVGOPPSQ*PEDSRS AFKQFF*NHTTPTDTSNEPDFVDPDKLWSDDN EEENVNDTLINGWYKNGKIPEHAPTYEFTFRRF FDDNGYFYNYPKPIEFSYNGQSEQQSDEDD QNTGSBIKNRDLVYV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPELPGFWF SCFSLPF*LGYRHCPFCANSFTTLPRLVLNTWA QAFFQPPPVLGLQV VGQAGLELLTSGDLFALASQSAGITGISHRAR PENGFENIF 519 1869 A 3994 751 126 NGGRHVGLCRTCLVNQMFASSILGKSHHBIS LISINQGHNALWKAAGPLPLKAGYCQSFSPC DSLKYQSWDEKDLTVPQRTHKRSVLRWIS QRGKLLAVEMEEGHCLLUPLGTECLGIKVPV		l	l			Í	
DVGRKSVTLRKMQEIKEMSPTPVELHKVT	j.	J]	J		ļ	
LYKDSDMEDFGFSVADGLLEKGYYVKNIRPA GPÖDLGGLKPYDRLLQVNHVRTRDFDCCLV	1			1			
SITE		1					
VPLIAESGNKLDLVISRNPLASOKSIDQOSLPG DYSEQNSAFFQQRSHGGNLETREPITNIL	1					i	LYKDSDMEDFGFSVADGLLEKGVYVKNIRPA
D*SEQNSAFFQQFSHGGNLETREPTINTL		1	[ĺ	GPGDLGGLKPYDRLLQVNHVRTRDFDCCLV
1864 A 3967 833 800		j					VPLIAESGNKLDLVISRNPLASQKSIDQQSLPG
NLGRSKSKQLFDYLIVIDFESTCWNDGKHHH SQEIIEFPAVLLNTSTIGQIDSEFQAYVQPQEHPI LSEFCMELTGIKQAQVDEGVPLKICLSQFCK WIHKIQQKNIIFATGISEPS/DF*SKIMCICYL VR*RISYTY*SKHKSKGC							
SQEIIEFPAVLLNTSTGQIDSEFQAYVQPQEHPI LSEFCMELTGIKQAQVDEGYPLKICLSQFCK WIHKIQQQKNIPATGISEPSDF*SKIMCICYL VR*RISYTY*SKHKSK.GC	514	1864	Α	3967	833	800	LEKQGVSGMATKRLARQLGLIRRKSIAPANG
LSEFCMELTGIKQQVDEGYPLKICLSQFCK WIHKIQQKNIIFATGISEPS/DF*SKIMCICYL VR*RISYTY*SKHKSKGC VR*RISYTY*SKHKSKGC CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL DLLGPEFLARKPLFKTKTYQSTF*SISKNE/FTC PNFIIEEGTDLIP*QVKHNPCHRLTPEGTVQL NRADS NRADS NRADS NRADS KMLC/QKESNYIRLKRAKMDKSMFVKIKTLGT GAFGEVCLARKVDTKALYATKTLRKKDVLL RNQVAHVKAERDILAEADNEWVVRLYYSFQ DKDNLYFVMDYIPGGDIMMSLLIRMGIFPESL ARFYIAELTCAVESVHKMGFHRDIKFDNILID RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAFEVLLRTGYTQL CDWSVGVILFEMLVGQPFFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIKLCRGFE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDPVDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV VFLDRVPLCHFGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRDISFTTLPRLVLNTWA QAFGPQPFKVLGLQV SPEMESHFITOGGVQWHILSSLQPLPFGFK*F SPEMESHFITOGGVQWHILSSLQPLPFGFK*F SPEMESHFITOGGVQWHILSSLQPLPFGFK*F SPEMESHFITOGGVQWHILSSLQPLPFGFK*F SPENGSPENTGAGVQWHILSSLQPLPGFK*F SPENGSPENTGAGVGQWHILSSLQPLPGGFK*F SPENGSPENTGAGVGQUFTAGAGVGQGVGGGCKAGVAGVAGVAGVQWHILSSLQPLPGGFK*F SPENGSPENTGAGVGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC		1	}	1			NLGRSKSKQLFDYLIVIDFESTCWNDGKHHH
LSEFCMELTGIKQQVDEGYPLKICLSQFCK WIHKIQQKNIIFATGISEPS/DF*SKIMCICYL VR*RISYTY*SKHKSKGC VR*RISYTY*SKHKSKGC CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL DLLGPEFLARKPLFKTKTYQSTF*SISKNE/FTC PNFIIEEGTDLIP*QVKHNPCHRLTPEGTVQL NRADS NRADS NRADS NRADS KMLC/QKESNYIRLKRAKMDKSMFVKIKTLGT GAFGEVCLARKVDTKALYATKTLRKKDVLL RNQVAHVKAERDILAEADNEWVVRLYYSFQ DKDNLYFVMDYIPGGDIMMSLLIRMGIFPESL ARFYIAELTCAVESVHKMGFHRDIKFDNILID RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAFEVLLRTGYTQL CDWSVGVILFEMLVGQPFFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIKLCRGFE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDPVDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV VFLDRVPLCHFGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRDISFTTLPRLVLNTWA QAFGPQPFKVLGLQV SPEMESHFITOGGVQWHILSSLQPLPFGFK*F SPEMESHFITOGGVQWHILSSLQPLPFGFK*F SPEMESHFITOGGVQWHILSSLQPLPFGFK*F SPEMESHFITOGGVQWHILSSLQPLPFGFK*F SPENGSPENTGAGVQWHILSSLQPLPGFK*F SPENGSPENTGAGVGQWHILSSLQPLPGGFK*F SPENGSPENTGAGVGQUFTAGAGVGQGVGGGCKAGVAGVAGVAGVQWHILSSLQPLPGGFK*F SPENGSPENTGAGVGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC	1	1	ĺ .				SQEIIEFPAVLLNTSTGQIDSEFQAYVQPQEHPI
VR*RISYTY*SKHKSKGC			'			1	LSEFCMELTGIKQAQVDEGVPLKICLSQFCK
1865 A 3969 492 182 CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL DLLGPEFLARKPLFKTKTYQSTF*SISKNEFTC PNFIIEEGTDLIF*QVKHNPCHRLTPEEGTVQL NRADS 516				i			
DLLGPEFLARKPLFKTRTYQSTF*SISKNE/FTC PNFIIEEGTDLIF*QVKHNPCHRLTPEEGTVQL NRADS 516 1866 A 3977 2 1357 KMLC/QKESNYIRLKRAKMDKSMFVKIKTLGI GAFGEVCLARKVDTKALYATKTLRKKDVLL RNQVAHVKAERDILAEADNEWVVRLYYSFQ DKDNLYFVMDYIPGGDMMSLLIRMGIFPESL ARFYIAELTCAVESVHKMGFHRDIKPDNILID RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPPYIAPEVLLRTGYTQL CDWWSVGVILFEMLVGQPPFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFF*NHTIPTDTSNFDPVDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRF FDDNGYPYYNYFRIEYEYINSQGSEQQSDEDD QNTGSEIKRNDLYYV 517 1867 A 3980 1358 1022 FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVTVNLPPSWD *RCRPPH/LANLCNFCRD\STTILPRLVLNTWA QAIFQPQPPFVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLF2*LGYRHVPPCLANSVFSVEMGVFLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAGYLPLKAGYC\QSFSPC DSLKYGISWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLLLPLGTECLGIK\PJV					•		VR*RISYTY*SKHKSKGC
PNFIEEGTDLIF*QVKHNPCHRLTPEEGTVQL NRADS	515	1865	Α	3969	492	182	CRFWGISTHCDTCDPLSPQTTEG**EGDLWSL
PNFIEEGTDLIF*QVKHNPCHRLTPEEGTVQL NRADS	ĺ	1					DLLGPEFLARKPLFKTKTYQSTF*SISKNE/FTC
NRADS							PNFIIEEGTDLIF*QVKHNPCHRLTPEEGTVQL
GAFGEVCLARKVDTKALYATKTLRKKDVLL RNQVAHVKAERDILAEADNEWVVRLYYSFQ DKDNLYFVMDYIPGGDMMSLLIRMGIFPESL ARFYIAELTCAVESVHKMGFHRDIKPDNILID RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGYILFEMLVGQPFFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE DRLGKNGADEIKAHPIFYNQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDP\VDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFIFRRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV QNFQPQPKVLGLQV GSFSPC SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLF2*LGYRHVPPCLANSVFSVEMGFLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF SINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV							
GAFGEVCLARKVDTKALYATKTLRKKDVLL RNQVAHVKAERDILAEADNEWVVRLYYSFQ DKDNLYFVMDYIPGGDMMSLLIRMGIFFESL ARFYIAELTCAVESVHKMGFIHRDIKPDNILID RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGVILFEMLVGQPPFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDP\VDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV 517 1867 A 3980 1358 1022 FFFKKFTQSLGFLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRDISFTTLPRLVLNTWA QAIFQPQPPKVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMGFLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYGSWDEKDLTYPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV	516	1866	A	3977	2	1357	KMLC/QKESNYIRLKRAKMDKSMFVKIKTLGI
RNQVAHVKAERDILAEADNEWVVRLYYSFQ DKDNLYFVMDYIPGGDMMSLLIRMGIFPESL ARFYIAELTCAVESVHKMGFIHRDIKPDNILID RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGVILFEMLVGQPPFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDP\VDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV 517 1867 A 3980 1358 1022 FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLFE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTYPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV]						
DKDNLYFVMDYIPGGDMMSLLIRMGIFPESL ARFYIAELTCAVESVHKMGFHRDIKPDNILID RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGVILFEMLVGQPPFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDPVDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV 517 1867 A 3980 1358 1022 FFFKKPTQSLGFLLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTLPRLVLNTWA QAIFQPQPPKVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV							
ARFYIAELTCAVESVHKMGFIHRDIKPDNILID RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGVILFEMLVGQPPFLAQITPLETQM KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDPVDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFIFRRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV 517 1867 A 3980 1358 1022 FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPKVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV							
RDGHIKLTDFGLCTGFRWTHDSKYYQSGDHP RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGVILFEMLVGQPPFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDP\VDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTFLPRLVLNTWA QAIFQPQPPKVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV			l i				
RQDSMDFSNEWGDPSSCRCGDRLKPLERRAA RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGVILFEMLVGQPPFLQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDPVDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV 517 1867 A 3980 1358 1022 FFFKKFTQSLGFLFSFSFFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHH\S LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTYPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV	1						
RQHQRCLAHSLVGTPNYIAPEVLLRTGYTQL CDWWSVGVILFEMLVGQPPFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDP\VDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV 517 1867 A 3980 1358 1022 FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV							
CDWWSVGVILFEMLVGQPPFLAQTPLETQM KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDPVDPDLWSDDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRFF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV 517 1867 A 3980 1358 1022 FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV							
KVINWQTSLHIPPQAKLSPEASDLIIKLCRGPE DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDP;VDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRFF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV 517 1867 A 3980 1358 1022 FFFKKFTQSLGFLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLFE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHH\S LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDT\KRSVLRWIS QRGK\LAVEMEEG\CL\L\LPLGTECLGIK\PIV				1	j		
DRLGKNGADEIKAHPIF*NQFDFSQ*PEDSRS AFKQFP*NHTTPTDTSNFDP\VDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRFF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV 517 1867 A 3980 1358 1022 FFFKKFTQSLGFLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHH\S LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDT\HKRSVLR\WIS QRGK\LAVEMEEG\HCL\\L\PLGTECLGIK\PIV					1		
AFKQFP*NHTTPTDTSNFDP\VDPDKLWSDDN EEENVNDTLNGWYKNGKHPEHAFYEFTFRF FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV 517 1867 A 3980 1358 1022 FFFKKFTQSLGFLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF NQGLRHVGLCRTCLVNQMFASSILGKSHH\S LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV							
517 1867 A 3980 1358 1022 FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTFLPRLVLNTWA QAIFQPQPPKVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTYPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV]						
FDDNGYPYNYPKPIEYEYINSQGSEQQSDEDD QNTGSEIKNRDLVYV]	· - , [•
517 1867 A 3980 1358 1022 FFFKKFTQSLGFLLFSFSFLFSCFFFFHFVLFCY VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV	j i			j		j	
S17					•		
VFLDRVPLCHPGWSAVVQSQVT/VNLPPSWD *RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF NQGLRHVGLCRTCLVNQMFASSILGKSHH\S LISINQGHNALWKAAG\PLPLKAG\PC\QSFSPC DSLK\YG\SWDEKDLTVPQRDT\KRSVLR\WIS QRGK\LAVEMEEG\HCLL\\LPLGTECLGIK\PIV	517	1867	Ā	3980	1358	1022	<u> </u>
*RCRPPH/LANLCNFCRD\SFTTLPRLVLNTWA QAIFQPQPPKVLGLQV 518 1868 A 3986 974 666 SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHH\HS LISINQGHNALWKAAG\PLPLKAG\PC\QSFSPC DSLK\PCYG\SWDEKDLT\PQRDT\HKRS\PC\LS\PC\PC\PC\PC\PC\PC\PC\PC\PC\PC\PC\PC\PC\]						
QAIFQPQPPKVLGLQV				1	ĺ	ĺ	
SPEMESHPITQAGVQWHHLSSLQPLPPGFK*F SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV	j j		- 1			1	
SCFSLPE*LGYRHVPPCLANSVFSVEMG\FLH VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV	518	1868	A	3986	974	666	
VGQAGLELLTSGDLPALASQSAGITG\SHRAR PENGFENIF 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV				3,00	-/-	V-00	
PENGFENIF 519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV				ł	İ		
519 1869 A 3994 751 126 NQGLRHVGLCRTCLVNQMFASSILGKSHHHS LISINQGHNALWKAAG\PLPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV		1	İ	İ	ľ	Ì	
LISINQGHNALWKAAGYPLKAGYC\QSFSPC DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV	510	1860	 _	200/	751	106	
DSLKYG\SWDEKDLTVPQRDTHKRSVLRWIS QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV	717	1007	^	J774	/31	120	
QRGK\LAVEMEEGHCLL\LPLGTECLGIK\PIV	1 1	1	- 1			Į	
				ł	1		
HLFSSEMGE\NRPMVG\ARHVYSNAALLSFTP	(I	į	- 1		1	l	
	L		اـــــا	1			HLFSSEMGE\NRPMVG\ARHVYSNAALLSFTP

CPO ID	Laron	1 14 24	SEQ	T 50 - 10 - 10 - 10 - 10 - 10 - 10 - 10 -	18 0	
SEQ ID NO: of	SEQ ID NO: of	Met hod	ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	"00	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ĺ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutarnine, R=Arginine, S=Serine,
	1	l]	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
		i		sequence		nucleotide insertion
		· · ·				LRCLGGEKHKSGLHARPVIVPSLELHYDMDSI
					<u> </u>	AHV\FADLLLIITLPSYYIPFC
520	1870	Ā	3999	882	698	QSFRLSLLSSWDYRHM*PRLANF*TVFFCRDR/
521	1871	A	4011	1346	1178	SLALLPRLVSNSWPQAILPPRPPKVLGLQT FFF*ETVSCSAS*AGVRSHDNSSLQPPSPG\SSN
52.	10/1	^	1011	1540	1170	PPTSASHVAGATGTHHHAWLLSV
522	1872	Α	4015	2	377	QGIALLTRMGESVKHVTGGYKLRTRPLEFAA
			İ			IGDYLDTFALKLGTIDRIAQRIIKEEIEYLVELR
	Ì					EYGPVYSTWSALEGELAEPLEGVSACIGNCST
	<u></u>				ľ	AL*ELTDDMTEDFLFVLREYILYSDSMK
523	1873	Α	4018	341	19	ERVIHNQIQQAQRSPHIFNARRSS/PRPNIVELP
		ĺ				KVKEVCKTSKS/GQVIYKGVSIRLRANFLAEP
		ĺ				L*NRREWDEAIKVLKEKQ\FLSKMVYPANLSF
						GNEGDITSFPAK
524	1874	Α.	4020	1067	743	FFLRWSL/DSVAQAGVKWCNLGSLQAPPPGF
						TPFSCLSLPSSWDYRHPPPRLAN*LTNFLCF**
						RQGFTVLARMVLIS*PHDLPASASQSAGITGL
525	1875		4001	701	201	SHCSWPTSSILS
323	16/3	Α	4021	781	351	QFRVIFFFLRRSHSVAQAGMQWHDHSLLQPL
						PPRLKQ/F/SHLSPPSIWDYRRVPPCLVNFSIFF
						VETGSCQPCLQLLGSSNPPASASQSAGIAGISH QGQPE*SFDIRFACVIAALRETFQCLCSASRVN
]			NKIINRPTHPVESSF
526	1876	Α	4024	80	341	TPSSTSRGTEEQQSSKMAWQRREEKEHLNVR
	1070	••	.021		341	RSSAEDGWKADKP/VDG*TPGEDHLPTPSPFQ
						LHIHSSESQLHHSVKSPPSLSFRLM
527	1877	Α	4026	593	230	DFYLYPERKKRGQMMTAVSLTTRPQESVAFE
						DVAVYFTTKEWAIMG\PAERALYRDVMLEN
			ļ	•		YGGCGPL*CHPTSKPALVFS\LEQGKESCFSPA
			<u> </u>			TGSSLSRNDWRAGWIGYLELRRYTYLS
528	1878	Α	4028	1160	242	GTSELLCIQRWNWGPAFPPRPGLALAPTLQLL
]			VEMGSAKSVPVTPARPPPHNKHLARVADPRS
- 1						PSAGILRTPIQVESSPQPGLPAGEQLEGLKHAQ
						DSDPRSPTLGIARTPMKTSSGDPPSPLVKQLSE
						VFETEDSKSNLPPEPVLPPEAPLSSELDLPLGT
			ł			QLSVEEQMPPWNQTEFPSKQVFSKEEARQPT
J				i		ETPVASQSSDKPSRDPETPRSS\GSMRNRWKP\
			1		ĺ	NSSKVL\GKSPLHPSCQDDNSPGTLTLRQGKA
-						AFKPLSENVSELK\EGA\ILGTGR\LLKTEGRA
529	1879	Α	4039	2	366	WEQGQD\HDKENQHFPLVES
-27	10,7	^	4037	-	סטכ	KDMVLIMEMQSMITMKCPQYL*E*RKIPDITK
j	1			ļ		CW*GCGSTGILIFC/WS*PL*KTI*QPR*FKQI*T ILTIIYSIM*EHTFHNAGV*LSDIYPRFMKGYV
	[ĺ	HTEICT*MFIAVLFVVVKTWKQF
530	1880	Ā	4057	358	3	LLEVNGNTIVTVFTKAQNKKNKGSRSILFKQL
		*	"""		-	RKYGSRINLLKSKHDKNICTENYKT*MKEIEA
ļ	1	į				/DTDKWKDILCSWIRRIHMKDILCSWIGRTHV
i	ļ					VKISILPKVNYRFYLISIKIIMAI
531	1881	A	4061	50	278	TQGTEEIYKISSCEWVQASFSTPLITLHDFKIY
j						HKATVIKMVWYWHRQ*KFSKN/RIESSEIEPH
l						IYDQFIFDKGEKIIQEKGNSFFNN/MCWKNWIF
	J		j	1	J	T*KR
532	1882	A	4069	19	368	NDLLENFKFWE*FKE*LENINGTVTEKETGGV
ŀ				1		YKELSSPKYSGTRQFYGQTISNFPGKIISMVY
	l	1]		KLFQNTE/TEGRHPISLYEFRITLITIPNKDNIYL
				<u> </u>]	QIWMPVSLMNIVTLKCPT
533	1883	Α	4076	1	355	PIRKFTKVAG*KSNTPK*LAFLHINNEQFENKI/
		- 1		ĺ	ľ	ITNUPFIIASKRIKYSGISLTKEMKDLYTETLLR
I	ı	- 1	1			KIKEDTNKWKDI/SCFWVGR/LNTVKMPK/VIC

NO: of NO: of No: of N	SEO ID	SEQ ID	Met	SEQ	Predicted	Deadles de de	1 4-1
nucleotide sequence unice olide correspondity unice of the correspondity of the correspondity of the correspondity of the correspondity of the correspondity of the correspondity of the correspondity of the correspondity of the correspondity of the correspondity of peptide sequence propriets and correspondity of peptide sequence propriets of the correspondity of peptide sequence propriets of peptide seq							Amino acid sequence (A=Alanine C=Cysteine,
Docation Docation			noa	1			
uence unce correspondi gla fin fin and mino acid residue of reptide sequence of reptid	1						
Union	1]				
mino soid residue of peptide sequence piptide sequence piptide sequence seq		uence	1				M=Methionine, N=Asparagine, P=Proline,
residue of poptide sequence Y=Tyrosine, X=Unknown, **Sign codon, *possible nucloside decine pyrosible nucloside pyrosible nucloside decine pyrosible nucloside decine pyrosible nucloside decine pyrosible nucloside decine pyrosible nucloside pyrosible nucloside decine pyrosible nucloside decine pyrosible nucloside decine pyrosible nucloside decine pyrosible nucloside pyrosible nucloside decine pyrosible nucloside decine pyrosible nucloside decine pyrosible nucloside pyrosible nucloside decine pyrosible nucloside decine pyrosible nucloside decine pyrosible nucloside decine pyrosible nucloside nucloside nucloside nucloside pyrosible nucloside nuclosited nu	uence		l	914			Q=Glutamine, R=Arginine, S=Serine,
Popsible nucleotide deletion, \(\text{Popsible} \)							T=Threonine, V=Valine, W=Tryptophan,
	1		l			sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	-	1	j				
1884		1	1		sequence]	
1884							IFNAIPIKMPMMCMAKIEKNSS
GAPGOTTKAQSML/PGSEKLRHLSTTSHOVL QTRLVDAAKALNL/HCHCIDFRQAFDMGR DLQTTPKREPTTRYKENEL/SELMINANKOE EMKDMIVETILINTMKEELLIDDATIMBERKDIV VPENGEPVGTREIKCERQQELISRINQAVA NKLISSVDYLRESPVGTLERCLQSLEKSQDVS VHITSINYLKQILNAAVIPEVTHSGSSVTRM LWEQIKQIQQITTWVSPPATILEWKRKVAQPA ESLSASKLAKSICGPRTRLINSHEPA-ASALRQ ESLSASKLAKSICGPRTRLINSHEPA-ASALRQ LEAGHSGRLEKTEDL-WLR-VRXDHAPPLARLS LESRSLQDVLLHREPKLQGLGRQGYQVVLVLCDNWGGHPFCALKSVVPPDEKHNDLALEF HYMRSLFKHERLVDLHGSVUDNYVGGGSSIA VLLIMBERHEDLYTICLAGGILLERTLQIALDV VEGIRFLHSQGL-VREDIKLGSVULDKQNRAKI TDLGFCKPAMMSGSVGTPHMAPELIFGK YDNSVDVYAFGILFWYICSGSVKLPEAFERCA SXDHLWNNVRGABFERLPYPDEECWQLME ACWDGDPLKRFLIGIVQPMLQGIMNRLCKS NESQPIRGLDST TALMPHEANYEEIPLKTDKDMDGFESQLEVRE FIRKTRICLISPILLAHHWALCDSKDCCKLSKD HFALAFILTQKLIKGGPPLVLTEREKSPSN ASLQKVTELTREPVCIERGGILWWITDISIWMK HFRKRIWLRA STANDARSHALLAHWALCDSKDCCKLSKD S36	534	1884	A	4088	3	1931	
			1		-	1	GAPGODTK AOSMI VEOSEVI PUI STESHOVI
DLQTPKRLEYTRKEENLYSEMMIANTMEEKDD TIMBERD	1	1	l		1	1	OTT I TO A A KAI AT VILOUGE DETRICA FOR A CO.
EMKDMIVETLNITMKEELIDATNMEEKDUT VPENGEPVÖTREIKCCIRQIQELISRINQAVA NKLISSVDYJRESFVOTILERCLQSLEKSQDVS VPITSNYLKQILNAAYHVEVTFIBGSSVTRM LWEQIRQIIQRITWVSPPATILEWKREVAQEAI ESLSASKLAKSICSQFRIRINSSHEAFAASLSC LEAGHGRILERTEDLWT.PRENDHAPPALARLS LESRSLQDVLLHRKPKLGQELGRGQYGYVIVC CDNWGGHFPCALKSVVPPDEKHWNDLALEF HYMR.SJ.PKHERL.VDLHGSVIDYNYGGGSSIA VLLIMERLHRDLYTGLKAGITLETRLQIALDV VEGIRFIHSQGLVHRDIKLNSVLJEKKPRAMASLILER VPINSVDVYAFGILFWYICGSSVLPEAFRCA SKDHLWNINVRGARPERLPVTDEECWLJMC YDNSVDVYAFGILFWYICGSSVLPEAFRCA SKDHLWNINVRGARPERLPVTDEECWLJMC ACWGODPLKRPLLGIVQPMLQGIMMRLCKS NSEQPNRGLDDST MAPPHEANVEEIFLKTDKDMGGFESGLEVRE IFLKTRGLPSTLLAHWALCDSKDCGKLSKD HFALAFHLINQKLIKGIDPFLVLTPEKISPSIN ASLQKVTELTRKPCUEFKGTLWRITDSIWMK HRAKRIWLRA SKOKVTELTRKPCUEFKGTLWRITDSIWMK HNRKRIWLRA SKOKVTELTRKPCUEFKGTLWRITDSIWMK HNRKRIWLRA SKOKVTELTRKPCUEFKGTLWRITDSIWMK HNRKRIWLRA SKOKVTELTRKPCUEFKGTLWRITDSIWMK HNRKRIWLRA SKOKVTELTRKPCUEFKGTLWRITDSIWMK HNRKRIWLRA SKOKVTELTRKPCUEFKGTLWRITDSIWMK HNRKRIWLRA SKOKVTELTRKPCUEFKGTLWRITDSIWMK HNRKRIWLRA SKOKVTELTRKPCUEFKGTLWRITDSIWMK HNRKRIWLRA SKOKVTELTRKPCUEFKGTLWRITDSIWMK HNRKRIWLRA SKOKVTELTRKPCUEFKGTLWRITDSIWMK HNRKRIWLRA SKOKVTELTRKPCUEFKGTLWRITDSIWMK HNRKRIWLRA SKOKVTELTRKPCUEFKGTLWRITDSIWMK HNRKRIWLRA SKOKVTELTRKPCUEFKGTLWRITDSIWMK HNRKRIWLRA SKOKVTELTRKPCUEFKGTLWRITDSIWMK SKOKTELTRKPCUEFKGTLWRITDSIWMK SKOKTELTRKPCUEFKGTLWRITDSIWMK SKOKTELTRKPCUEFKGTLWRITDSIWMK SKOKTELTRKPCUEFKGTLWRITDSIWMK SKOKTELTRKPCUEFKGTLWRITDSIWMK SKOKTELTRKPCUEFKGTLWRITDSINGSTAGN SKORTENT SKORTEN		ĺ	l		· ·	1	DI OFFICEI EVTREVENEI VECI A GLADIVO
VPENGEPYGTREIKCCIRQIQELISRINOAD NKLISSYDYLRESYGTLERC(SQILESRODVS NKLISSYDYLRESYGTLERC(SQILESRODVS NKLISSYDYLRESYGTLERC(SQILESRODVS NKLISSYDYLRESYGTLERC(SQILESRODVS NKLISYDYLRESYGTLERC(SQILESRODVS NKLISYDYLRESYGTLERC(SQILESRODVS NKLISYDYLRESYGTLERC(SQILESRODVS NKRISYPERELT SQILEST NYMESPHERILRYNKDHAPRLARS LEARSLQVLLHRKPKLGGELGRGGYGYVYVL CDNWGGHFCALKSVVPPDEKHWNDLALE HYMRSLPKHERLYDHROSYNYGGGSIA VLLIMERLHRDLYTGLKAGITLETRLQIALDV VEGIRFLHSGILVHRDILKTYNYGGGSIA VLLIMERLHRDLYTGLKAGITLETRLQIALDV VEGIRFLHSGILVHRIDKONPAKI TDLGFCKPEAMMSGSIGTPHMAPELFFGK YDNSVDVYAFGILFWYICSGSVKLPEAFGK YDNSVDVYAFGILFWYICSGSVKLPEAFGK YDNSVDVYAFGILFWYICSGSVKLPEAFGK YDNSVDVYAFGILFWYICSGSVKLPEAFGK YDNSVDVYAFGILFWYICSGSVKLPEAFGK YDNSVDVYAFGILFWYICSGVKLPEAFGK YDNSVDVYAFGILFWYICSGVKLPEAFGK YDNSVDVYAFGILFWYICSGVKLPEAFGK YDNSVDVYAFGILFWHOPELSYDDEN ALMPHEANYEEIFLKTÜMMOFFESGIEVE FILKTRGLPFTLAHIWALCDSKOCKLSVB HFALAPHLITQKLIGGIDPPLVLTPEKISPSNR ASLQKVTELTREVCIFGTIL WRITDSIWMK HNRKRIWLRA TSLGT T		ļ	!				ENVENTED TRAKENEL TESLIVINIANKAVE
NKLISSYDYLRESPYGTLERCLQSLEKSQOYS			İ				EMADIVI VETENTIMAEELLUDA I NMEFKUVI
VHITSYYLKQILNAAYHVEVTFHISOSYTRM				'	1		VPENGEPVGTREIKCCIRQIQELIISRLNQAVA
LWEQIKQIIQRITWVSPPAITLEWKRKVAQEAL ESLSASKLAKSICSQFRTRLISSHEAFAASLRQ LEAGHSGRLEKTEDL.WLRVRKDHAPRLARLS LESRSLQDVLLHRIKPKLGQELGRGQYQVYVL CDNWGGHFPCALKSVVYPDEKHWNDLALEF HYMRSLFKHERLJOLHGSVIDJYNYGGGSIA VLLIMERLIFRILYTICKAGILTETRLQIALDV VEGIRFLHSQGLVHRDIKLKNVLLDKQNRAKI TDLGFCKPEAMMSGSIVGTPHIMAPELFTGK YDNSUDVYAFGILFFYYGCSGSKIA VLLIMERLIFRILYTICKAGILFYPTDEECWQLME ACWDGPLKRPLGIVQPMLQGIMRLCKSI NSEQPNRGLDDST 1885 A 4090 2 417 ALMPHEANYBEIFLKTDKDMGFFSGLEVRE IFLKTRGJESFILAHWALCDSKDCGKLSKD HFALAFHLITIQKLIKGDPPLVLTPEKISPSNR ASLQKVTELTREPVCIFKGTILWRITDSIWMK HNRKRIWLRA 536 1886 A 4102 569 829 DHQK*KNIPCSWIGRINIVKMSILPKAIYFSAI PIKIPMTFFIE*SNVYRITTKTQE*AKAILSKK EQNLEESHYLDFK*YYRAV 537 1887 A 4104 54 281 SIDCHILIRRULJDFSKRITIAQIKEHKWML IEVPVQRPVLYPQEQENEPSIGEFNEQVLRLM HSLGGDQKTIE 538 1888 A 4109 141 314 RRHIPLKIRSVVSHLKCFYKFILTFFFAGCSQPL VPRENITAWMANAIGLITALPVS 539 1889 A 4111 268 1 ASRPWGHSY*PFNQQEDYTLKRPIASSEI*MM PLKFATKKSPGPYBPTAEFSHTKEDLVPILW PLFPKIYREGTLPHSFYESHTLKEDLVPILW PLFPKIYREGTLPHSFYESHTKEDLVPILW PLFPKIYREGTLPHSFYESHTKEDLVPILW PLFPKIYREGTLPHSFYESHTKEDLVPILW PLFPKIYREGTLPHSFYESHTLAGHT MLLKYTYGRARTITDEKEGVLLDFILSAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLLTCMCCMFMDRNAQEVLSSGGFLSCAAE AALGDFLGLHRRTQQPAVDRLLSDSAQWR VRGHGVMESGRAPQPOGRRGRRPRKPRP GRWRREGCGAGGROVCVAAWSQRSLAGNN DYRLFHKMSNSHPLRPFTAVGEIDLYBLSH IGALLIGEEVGDVTFVVEKKRPAHRVILABH IGALLIGEEVGDVTFVVEKKRPAHRVILABH IGALLIGEEVGDVTFVVEKKRPAHRVILABH IGALLIGEEVGDVTFVVEKKRPAHRVILABH IGALLIGEEVGDVTFVVEKKRPAHRVILABH IGALLIGEEVGDVTFVVEKKRPAHRVILABH IGALLIGEEVGDVTFVVEKKRPAHRVILABH IGALLIGEEVGDVTFVVEKKRPAHRVILABH IGALLIGEEVGDVTFVVEKKRPAHRVILABH IGALLIGEEVGDVTFVVEKKRPAHRVILABH IGALLIGEEVGDVTFVVEKKRPAHRVILABH IGALLIGEERGERGLAGOFSINHVRILLHWDR SSRSYSYFIEVSMDELDWVRVIDHSQVTLCRS VGRYRGERGERGLAGOFSINHVRILLHWDR DSRSYSYFIEVSMDELDWVRVIDHSQVTLCRS VGRYTKATLTICKGLIVPNENVATIADCASVI ECMFTKKTFTLEKGLIVPNENVATIADCASVI ECMFTKKTFTLEKGLIVPNENVATIADCASVI ECMFTKKTFTLEKGLIVPNENVATIADCASVI ECMFTKKTFTLEKGLIVPNENVATIADCASVI	ł	1	l	!			
ESLSASKLAKSICSORTRILNSSHEARASLRQ LEAGHSGRILKTEDL WARVROHAPRIARIZ LESRSLQDVLLHRKPKLGQELGRGQYGVVVL CDNWGGHFFCALSSVVPPDEKHWNDLAEF HYMRSLKKERLJUDHGSVUDYNYGGGSSLA VLLMRELJRDLYTGLKAGI.TLETRLQJALDV VEGIRLHSQG.VHRDIKKINVLLDKQNRAKI TDLGFCKPEAMMSGSIVGTPIHMAPELFTGK YDNSVDVVAFGILFWYICSGSVKLPEAFERCA SKDHLWNNVRRGARPERLPVPDEECWQLME ACWDGDPLKRPLLGIVQPMLQGIMNRLCKSI NSEQPNRGLDDST 335 1885 A 4090 2 417 ALMPHEANYEEIFLKTDKDMDGFSEGGLEVRE IFLKTRGJPSTLAHHWALGSKEGGKLSKD HFALAFHLIT\QKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTREPVCIIFKGTILWRITDSIWMK HNRKIWLRA 536 1886 A 4102 569 829 DHQK*KNIPCSWIGKNIVKMSILPKATYRFSAI PIKIPMTFFTE**SNVYRTTKTQ**AKAILSKK EQNLESSHYLDFK*YYRAV 537 1887 A 4104 54 281 SIDCEHLIRRMLVLDPSKRLTIAQIKEHKWML IEVPVQRPVLYPQEQENEPSIGEFNEQVLRLM HSLGDQQKTE IEVPVQRPVLYPQEQENEPSIGEFNEQVLRLM HSLGDQQKTE ST ST ST ST ST ST ST ST ST ST ST ST ST				1			
LEAGHSGRLEKTEDLWR.VRV.DH.PRI.ARIS. LESRS.LOVD.LHR.PRYL.GGL.GR.GQVGVVV.L CDNWGGHFPCALKSVVPPDEKHWNDLALEF HYMRSL.PKHERL.VDLHGSVIDYNYGGSSIA. VLLIMER.LHRDLYTGL.KGGLT.LETRL.QIALDV VEGIRFLHS.OGL.VHRDIKLK.NVLLDK.ORT.ARI TDLGFCK.PEAMMSGSIV.GTPHIMAPELFTGK YDNSVDVYAFGILFWYICSGSVKLPEAFERCA SKDHLWNNVRGARFER.PLYPUECEWQLME ACWDGDPLK.RPLLGIV.OPM.QGIMNRL.CKS NSEQPNGLDDST 335 1885 A 4090 2 417 ALMPHEANYEEFLK.TDK.DMDGFESGL.EVRE IFLKTR.GL.PSTLLAHIWALCDSK.DCGKL.SKD HFALAFHLTTQKLIKGID.PPL.VLTPEKISPSIN ASLQKVTELTREPVCUIFKGTILWRITDSIWMK HNRKRIWLRA 536 1886 A 4102 569 829 DHGK*KINFCSWIGRINVKMSILPKATYRFSAI PIKLPMTFFTEI*S*NVYRTTKTQE*AKAILSKK EQNLEESHYLDFK*YYRAV 357 1887 A 4104 54 281 SIDCEHLIRR.ML.VLDPSKR.TIAQIKEHKWML IEVPVQRPVL.YPQCQENEPSIGEFNEQVLRLM HSLGIDQOKITE 358 1888 A 4109 141 314 IRHIPLKIRSVVSHLKCFYKFILTFFFAGCSQPL VPRENITAWMANIGLITTALPVS 539 1889 A 4111 268 1 ASRPWGBSYP*PNQQEVDTIK.KPIASSEI*MM I*KFATNKS.SPGYPRTAEFSHTFKEDL.VPILW PLFPKLYREGTLPISFYEAPESHTFKEDL.VPILW PLFRLYREGTLPISFYEAPESHTFT			l			1	LWEQIKQIIQRITWVSPPAITLEWKRKVAQEAI
LEAGHSGRLEKTEDLWR.VRV.DH.PRI.ARIS. LESRS.LOVD.LHR.PRYL.GGL.GR.GQVGVVV.L CDNWGGHFPCALKSVVPPDEKHWNDLALEF HYMRSL.PKHERL.VDLHGSVIDYNYGGSSIA. VLLIMER.LHRDLYTGL.KGGLT.LETRL.QIALDV VEGIRFLHS.OGL.VHRDIKLK.NVLLDK.ORT.ARI TDLGFCK.PEAMMSGSIV.GTPHIMAPELFTGK YDNSVDVYAFGILFWYICSGSVKLPEAFERCA SKDHLWNNVRGARFER.PLYPUECEWQLME ACWDGDPLK.RPLLGIV.OPM.QGIMNRL.CKS NSEQPNGLDDST 335 1885 A 4090 2 417 ALMPHEANYEEFLK.TDK.DMDGFESGL.EVRE IFLKTR.GL.PSTLLAHIWALCDSK.DCGKL.SKD HFALAFHLTTQKLIKGID.PPL.VLTPEKISPSIN ASLQKVTELTREPVCUIFKGTILWRITDSIWMK HNRKRIWLRA 536 1886 A 4102 569 829 DHGK*KINFCSWIGRINVKMSILPKATYRFSAI PIKLPMTFFTEI*S*NVYRTTKTQE*AKAILSKK EQNLEESHYLDFK*YYRAV 357 1887 A 4104 54 281 SIDCEHLIRR.ML.VLDPSKR.TIAQIKEHKWML IEVPVQRPVL.YPQCQENEPSIGEFNEQVLRLM HSLGIDQOKITE 358 1888 A 4109 141 314 IRHIPLKIRSVVSHLKCFYKFILTFFFAGCSQPL VPRENITAWMANIGLITTALPVS 539 1889 A 4111 268 1 ASRPWGBSYP*PNQQEVDTIK.KPIASSEI*MM I*KFATNKS.SPGYPRTAEFSHTFKEDL.VPILW PLFPKLYREGTLPISFYEAPESHTFKEDL.VPILW PLFRLYREGTLPISFYEAPESHTFT							ESLSASKLAKSICSQFRTRLNSSHEAFAASLRQ
LESSILQDVILHRRPKI.GQELGRQGYGVVYL CDNWGGIFPCALKSVYPPDEKHWNDLALEF HYMRSLPKHERLVDLHGSVIDYNYGGGSSIA VILIMERLHRDLYTGLK.AGI.TLETRLQIALDV VEGIRFLIRSQCI.HYRDIK.NVILDIKQNRAKI TDLGFCKPEAMMSGSIVGTPIHMAPELFTGK YDNSVDVYAPGILFWYICSGSVKLPEAFECA SKDHLWNNYRGARPERLPVPDEECWQLME ACWDGDPLKRPLLGIVQPMLQGIMNRLCKS NSEQPNRGLDDST ALMPHEANYEEIPLKTDK.DMDGFESGL.EVRE IFLKTR.GLPSTIL.AHIWALCSGISCKLSKAL HFALAFHLITQKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTREPVCIIFKGTILWRITDSIWMK HRALAFHLITQKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTREPVCIIFKGTILWRITDSIWMK HNKRRIWLRA 536	1		ŀ				LEAGHSGRLEKTEDLWLRVRKDHAPRLARLS
CDNWGGHFPCALKSVVPPDEKHNDLALEF	1		ł			}	LESRSLODVLLHRKPKLGOELGRGOYGVVYL
HYMRSLPKHERLVDLHGSVIDTNYGGGSSLA VLLMBELHRDLYTGLKAGLTLETRLQIALDV VEGIRFLHRQLVTGLKAGLTLETRLQIALDV VEGIRFLHSQLVHRDIKLKNVLLDKQNRAKI TDLGFCKPEAMMSGSIVGTPHMAPELFTGK YDNSVDYAPGILFWYICSGSVKLPEAFERCA SKDHLWNNVRGARRPERLPVFDEECWQLME ACWDGDPLKRPLLGIVQPMLQGIMNRLCKS NSEQPNRGLDDST ALMPHEANYEEIFLKTDKDMDGFESGLEVKD HEALAFHLITQKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIJFKGTLWRITDSIWMK HFALAFHLITQKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIJFKGTLWRITDSIWMK HFALAFHLITQKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIJFKGTLWRITDSIWMK HFALAFHLITQKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIJFKGTLWRITDSIWMK HFALAFHLITQKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIJFKGTLWRITDSIWMK HFALAFHLITQKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIJFKGTLWRITDSIWMK HFALAFHLITQKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIJFKGTLWRITDSIWMK HFALAFHLITQKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIJFKGTLWRITDSIWMK HFALAFHLITQKEHKWML EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVQRVLTPGEONEPSIGEFNEQVLR EVPTQRVTTAWNAIGHTALTVT EVPTQRVTTAWNAIGHTALTVT EVPTQRVTTAWNAIGHTALTVT EVPTQRVTTAWNAIGHTALTVT EVPTQRVTTAWNAIGHTALTVT EVPTQRVTTAWNAIGHTALTVT EVPTQRVTTAWNAIGHTALTVT EVPTQRVTTAWNAIGHTALTVT EVPTQRVTTAWNAIGHTALTVT EVPTQRVTTAWNAIGHTAMNAIGHTA EVPTQRVTT EVPTQRVTTAWNAIGHTAM EVPTQRVTT EVPTQRVT	1						
VILIMERIHRDI, YTGI, KAGITI, ETRIQIALDY VEGIRFHSQGI, VHRDIKI, KIVILDK, QNRAKI TDLGFCKPEAMMSGI, VTRIPHMAPEL, FTGK YDNSVDVYAPGII, FWYICSGSVKLPEAPERCA SKDHLWINVRRGARER, PVPEECWQI, ME ACWDGDPLKRPLLGIVQPMLQGIMNRLCKS NSEQPINGLDDST ALMPHEAYYEEIFLKTDKDMDGFESGLEVRE IFLKTRGLPSTLLAHIWALCDSKDCGKLSKD HFALAFHLIYKLIKGDPPLVLTPEKISPSNR ASLQKVTELTREPVCLIFKGTLWRITDSIWMK HNRKRIWLRA ASLQKVTELTREPVCLIFKGTLWRITDSIWMK HNRKRIWLRA ASLQKVTELTREPVCLIFKGTLWRITDSIWMK HNRKRIWLRA DHQK*KNIPCSWIGRNIVKMSILPKAIYRFSAI PIKIPMTFTEI*S*NVYRTTKTQP*AKAILSKK EQNLEESHYLDFK*YYRAV SIDCEHLRRMIVLYDPSKRLTIAQIKEHKWML IEVPVQRPVLYPQEQENEPSIGEFNEQVLRLM HSLGIDQQKTTE STORE THE STO							
VEGIRFLHSQGLVHRDIKLKNYLLDKQNRAKI TDLGFCKPAMMSGSIVGTPHMAPELFTGK YDNSVDVYAFGILFYPTGEKCWQLME ACWDGDPLKRPLIGIVQPMLQGIMNRLCKS\ NSEQPNRGLDDST ALMPHEANYEEJFLKTDKDMDGFESGLEVRE FILKTRGLPSTLLAHIWALCDSKDCGKLSKD HFALAFHLITVOKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIJFKGTILWRIDSIWMK HNRKRIWLRA ASLQKVTELTRKPVCIJFKGTILWRIDSIWMK HNRKRIWLRA ASLQKVTELTRKPVCIJFKGTILWRIDSIWMK HNRKRIWLRA ASLQKVTELTRKPVCIJFKGTILWRIDSIWMK HNRKRIWLRA ASLQKVTELTRKPVCIJFKGTILWRIDSIWMK HNRKRIWLRA ASLQKVTELTRKPVCIJFKGTILWRIDSIWMK HNRKRIWLRA ASLQKVTELTRKPVCIJFKGTILWRIDSIWMK HNRKRIWLRA SIDCHHLIRMILVJDPSKRITIAQIKEHKWML HEVPVQRPVLYPQEQENEPSIGEPNEQVLRLM HSLGIDQQKTIE EVPVQRPVLYPQEQENEPSIGEPNEQVLRLM HSLGIDQQKTIE STAPPENTAWMNAIGLIITALPVS ASSPWGHSYP*FNQQEVDTLKRPIASSEI*MM PRFATKKSPGPYRFTAEPSHTFKEDLVPILW PLFPKIYREGTLPHSFYEASITL FSPCAGRAATFWGPLFWRGRGSGRCEKAAE AALGDFLGHRRTQQPAVDRLISDASAQWR VRGHGGVRESGRAPQQPGRRGRRPRKRPR GRWREGCGAGGRGVCVAAWSQRSIAGNN DYRL*HKMSNISHPLRPFTAVGEIDHVHILSEH IGALLIGGEYGDVTFVVEKKRFPARVILAAR CQYFRALLYGGMRESQPEADPLQDTTAEAFT MLLKYTYGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLITCMCCMFMDRNAQEVLSSEGFLSLSK TALLINIVCRIPSKAPAPEDDFLALLTMVCKHNSK ENHAELMQAVRLPLMSLTELLNIVCKHNSK ENHAELMQAVRLPLMSLTELNIVTNKIFHVAF ECMTINKTITLKGLIVPMENVATLADCASVI ECMTINKTITLKGLIVPMENVATLADCASVI ECMTINKTITLKGLIVPMENVATLADCASVI	ļ j		ļ		,		
TDLGFCKPEAMMSGIVGTPHMAPELFTGK							VEGIDEI HOOGI VUDDIKI KARI I DVONDAKI
YDNSVDVYAFGILFWYICSGSVKLPEAFERCA SKDHLWNIVRGARFERLPVFDEECWQLME ACWDGDPLKRPLLGIVQPMLQGIMRLCKS NSEQPNRGLDDST	l. l						
SKDHLWNNVRGARPERLPYFDEECWQLME ACWDGPLKRPLLGIVQPMLQGIMNRLCKSI NSEQPNRGLDDST 1885 A 4090 2 417 ALMPHEANYEEIFLKTDKDMDGFESGLEVRE IFLKTRGLPSTLLAHIWALLOSKDCGKLSKD HFALAFHLITQKLIKGDPPLVLTTEKISPSNR ASLQKVTELTRKPVCIIFKGTILWRITDSIWMK HNRKRIWLRA 536 1886 A 4102 569 829 DHQK*KNIPCSWIGRINIVKMSILPKAIYRFSAI PIKIPMTFFTEI*S*NVYRTTKTQE*AKAILSKK EQNLEESHYLDFK*YYRAV 537 1887 A 4104 54 281 SIDCEHLIRRMLVLDPSKRLTIAQIKEHKWML IEVPVQRPVLYPQEQENEPSIGEFNEQVLRLM HSLGIDQQKTIE 1888 A 4109 141 314 IRHIPLKIRSVVSHLKCFYKFILTFFFAGCSQPL VPRENITAWMNAIGLIITALPVS 539 1889 A 4111 268 1 ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM I*KFATKKSSPGPYRFTAFESHTFKEDLVPILW PLFFKIYREGTLPHSFYEASITL 540 1890 A 4142 198 2064 PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGHRRTQQPAVDRLLSDASAQWR VRCHGGVRESGRAPQQPGRRRGRRPRKRRR GRWRREGGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTTVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTISFLICTIINIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALINVCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAALDAIKVRSESRDMDLMYRGMLDEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLQQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFFLEEGLIVPMENVATIADCASYI							
ACWDGPPLKRPLLGIVQPMLQGIMNRLCKS S35							
NSEOPNRGLDDST	, ,						
1885				•			
IFLKTR/GLPSTLLAHIWALCDSKDCGKLSKD HFALAFHLIT/QKLIKGDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIIFKGTILWRITDSIWMK HRKRIWLRA 536							
HFALAPHLITOKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIIFKGTILWRITDSIWMK ASLQKVTELTRKPVCIIFKGTILWRITDSIWMK HNRKRIWLRA	535	1885	Α	4090	2	417	
HFALAPHLITOKLIKGIDPPLVLTPEKISPSNR ASLQKVTELTRKPVCIIFKGTILWRITDSIWMK ASLQKVTELTRKPVCIIFKGTILWRITDSIWMK HNRKRIWLRA	}						IFLKTR/GLPSTLLAHIWALCDSKDCGKLSKD
ASLQKYTELTRRPVCIIFKGTILWRITDSIWMK HNRKRIWLRA	,				·		
HNRKRIWLRA							
1886 A 4102 569 829 DHQK*KNIPCSWIGRINIVKMSILPKAIYRFSAI PIKLIPMTFFTEI*\$\sqrt{1}\s					,		
PIKIPMTFFTEI*S*NVYRTTKTQE*AKAILSKK EQNLESSHYLDFK*YYRAV 537 1887 A 4104 54 281 SIDCEHLIRR/MVLDPSKRLTIAQIKEHKWML EVPVORPVLYPQEQENEPSIGEFNEQVLRLM HSLGIDQQKTIE EVPVORPVLYPQEQENEPSIGEFNEQVLRLM HSLGIDQQKTIE 538 1888 A 4109 141 314 RHIPLKIRSVVSHLKCFYKFILTFFFAGCSQPL VPRENITAWMAIGLIITALPVS 539 1889 A 4111 268 1 ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM I*KFAT\KKSPGPYRFTAEFSHTFKEDLVPILW PLFPKIYREGTLPHSFYEASITL 540 1890 A 4142 198 2064 PEPGAGRAATPWGPLFWGRGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRRGRPRKRPR GRWREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVITVVEKKRPPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVILDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI	536	1886	A	4102	569	829	
537 1887 A 4104 54 281 SIDCEHLIRRILVILDPSKRLTIAQIKEHKWML IEVPVQRPVLYPQEQENEPSIGEFNEQVLRIM HSLGIDQQKTIE 538 1888 A 4109 141 314 IRHIPLKIRSVVSHLKCFYKFILTFFFAGCSQPL VPRENITAWMAIGLIITALPVS 539 1889 A 4111 268 I ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM I*KFATKKSPGPYRFTAEFSHTFKEDLVPILW PLFPKYYREGTLPHSYYEASITL 540 1890 A 4142 198 2064 PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGOFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVVRSGRAPQQPGRRRGRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TAALLNIVLRDSFAAPEKDIFLALLNIVCKHNSK ENHAELMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSINHVRILLWDR DSRSYSYPIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTINKTFTLEKGLIVPMENVATIADCASVI					•		
1887 A 4104 54 281 SIDCEHLIRRVILVLDPSKRLTIAQIKEHKWML IEVPVQRPVLYPQEQENEPSIGEFNEQVLRIM HSLGIDQQKTIE	j ,						
ievpvqrpvlypqeqenepsigefneqvlrlm hslgidqqktie ibvpvqrpvlypqeqenepsigefneqvlrlm hslgidqqktie irhiplkirsvvshlkcfykfiltfffagsqpl vprenitawmnaigliitalpvs i asrpwghsyp*fnqqevdtlkrplassei*mm i*kfatkkspqpyrftaefshtfkedlvpilw plfpkiyregtlphsfyeasitl i asrpwghsyp*fnqqevdrlksdasqqwr plfpkiyregtlphsfyeasitl i*kfatkkspqpyrftaefshtfkedlvpilw plfpkiyregtlphsfyeasitl i*kfatkkspqpyrftaefshtfkedlvpilw plfpkiyregtlphsfyeasitl i*kfatkkspqpyrftaefshtfkedlvpilw plfpkiyregtlphsfyeasitl i*kfatkkspqpyrftaefshtfkedlvpilw plfpkiyregtlphsfyeasitl i*kfatkkspqpyrftaefshtfkedlvpilw plfpkiyregtlphsfyeasitl i*kfatkkspqpyrftaefshtfkedlvpilw plfpkiyregtlphsfyeasitl i*kfatkkspqpyrftaefshtmilkspqrggregtlphsfyeasitl plfpkiyregtlphsfyeasitl i*kfatkkspqpyrftaefshtmilkspqrggregtlphsfyeasitl plfpkiyregtlphsfyeasitl i*kfatkkspqpyrftaefshtmilkspqrggregtlphsfyeasitl plfpkiyregtlphsfyeasitl i*kfatkkspqpyrftaefshtmilkspqrggregtlphsfyeasitl plfpkiyregtlphsfyeasitl i*kfatkkspqpyrftaefshtmilkspqrggregtlphsfyeasitl plfpkiyregtlphsfyeasitl i*kfatkkspqpyrftaefshtmilkspqrggregtlphsfyeasitl plfpkiyregtlphsfyeasitl i*kfatkkspqpyrftaefshtmilkspqrggregtlphsfyeasitl plfpkiyregtlphsfyeasitl i*kfatkkspqpyrftaefshtmilkspqrggregtlphsfyeasitl plfpkiyregtlphsfyeasitl plfpki	537	1887	Δ.	4104	54	201	
HSLGIDQQKTIE	33,	100,	Λ	4104	24	201	
1888 A 4109 141 314 IRHIPLKIRSVVSHLKCFYKFILTFFFAGCSQPL VPRENITAWMNAIGLIITALPVS 1889 A 4111 268 1 ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM I*KFATKKSPGPYRFTAEFSHTFKEDLVPILW PLFPKIYREGTLPHSFYEASITL 1890 A 4142 198 2064 PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQOPGRRGRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVIFVVEKKFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI							
VPRENITAWMNAIGLITALPVS 539 1889 A 4111 268 1 ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM I*KFATKKSPGPYRFTAEFSHTFKEDLVPILW PLFPKIYREGTLPHSFYEASITL 540 1890 A 4142 198 2064 PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRGRPRKRPR GRWRREGCGAGGRGVCVAAWSORSLAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVIFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAELMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI	520	1000		4100			
1889 A 4111 268 I ASRPWGHSYP*FNQQEVDTLKRPIASSEI*MM I*KFAT\KKSPGPYRFTAEFSHTFKEDLVPILW PLFPKIYREGTLPHSFYEASITL 1890 A 4142 198 2064 PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRRGRRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGFHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI	338	1888	A	4109	141	314	
I *KFATKKSPGPYRFTAEFSHTFKEDLVPILW PLFPKIYREGTLPHSFYEASITL 540 1890 A 4142 198 2064 PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRRGRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLITDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI							
PLFPKIYREGTLPHSFYEASITL 540 1890 A 4142 198 2064 PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGLHRRTQOPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRRGRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI	539	1889	A	4111	268	1	
1890 A 4142 198 2064 PEPGAGRAATPWGPLFWRGRGSGRCEKAAE AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRRGRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI]	i					I*KFAT\KKSPGPYRFTAEFSHTFKEDLVPILW
AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRRGRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI				[
AALGDFLGLHRRTQQPAVDRLLSDASAQWR VRGHGGVRESGRAPQQPGRRRGRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI	540	1890	A	4142	198	2064	
VRGHOGVRESGRAPQOPGRRRGRPRKRPR GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI							
GRWRREGCGAGGRGVCVAAWSQRSIAGNN DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI							
DYRLFHKMSNSHPLRPFTAVGEIDHVHILSEH IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI							
IGALLIGEEYGDVTFVVEKKRFPAHRVILAAR CQYFRAILYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLITDEKEEVILDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLIGQPSINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI					ľ	l	
CQYFRALLYGGMRESQPEAEIPLQDTTAEAFT MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI		-	ĺ		-		
MLLKYIYTGRATLTDEKEEVLLDFLSLAHKY GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI		1		I	1	Į.	
GFPELEDSTSEYLCTILNIQNVCMTFDVASLY SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI	' ·	1	l	l		ļ	
SLPKLTCMCCMFMDRNAQEVLSSEGFLSLSK TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI						ļ	
TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI				ŀ			
ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI				1			
DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI				İ		l	TALLNIVLRDSFAAPEKDIFLALLNWCKHNSK
DAILDAIKVRSESRDMDLNYRGMLIPEENIAT MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI	ŀ	1	Į	1	1	Į.	ENHAEIMQAVRLPLMSLTELLNVVRPSGLLSP
MKYGAQVVKGELKSALLDGDTQNYDLDHG FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI		İ		1	}		DAILDAIKVRSESRDMDLNYRGMLIPEENIAT
FSRHPIDDDCRSGIEIKLGQPSIINHVRILLWDR DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI	ĺ		1	l	l		MKYGAQVVKGELKSALLDGDTQNYDLDHG
DSRSYSYFIEVSMDELDWVRVIDHSQYLCRS WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI	i		1	l	[FSRHPIDDDCRSGIEIKI.GOPSIINHVRII I WOD
WQKLYFPARVCRYIRIVGTHNTVNKIFHIVAF ECMFTNKTFTLEKGLIVPMENVATIADCASVI		}		į			DSRSYSYFIEVSMDFI DWVRVIDHSOVI CDS
ECMFTNKTFILEKGLIVPMENVATIADCASVI	}	ł	ĺ	- 1	l l		
ECUIT INTITILE ROLLYTIMEN VAIIAD CASVI	}	į		ŀ	1		
	İ	ļ		ľ	1	1	ECVEDEDNAL INCOTENTIAL VALIADICASVI
EGVSRSRNALLNGDTKNYDWDSGYTCHQLG	ļ						
SGAIVVQLAQPYMIGSIRVLLWDCDDRSY		1001		-,,,,	200		SGATY VQLAQPYMIGSTRVLLWDCDDRSY
541 1891 A 4146 282 778 GTLGYPNGARGQPQDNFFAHQ\VSHHPPISAC	J41	ולפו }	A	4146	282	778	GILGYPNGARGQPQDNFFAHQ\VSHHPPISAC

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	P=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence]	J	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1]	amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ł		l		peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
		1		sequence	Ì	nucleotide insertion
 	 	 	 	Sequence		HAESENFAFWQDMKWKNKFWGKSLEIVPVG
1		1	i			TVNVSLPRFGDHFEWNKVTSCIHNVLSGQRW
						IEHYGEVLIRNTQDSSCHCKITFCKAKYWSSN
				İ		VHEVQGAVLSRSGRVLHRLFGKWHEGLYRG
L						PTPGGQCIWKP
542	1892	Α	4147	44	433	SVDAYVCNDIVFSYRTTITLLEGA*LTHRYVA
					Ì	QDPKQGQLRSLHLTCDSAPAGSQGTWSTSCR
		1				INHLIFRGGAQITFLATFDDSPKAVLGDRLLLT
						ANVSSENNTPRTSKTTFQLELSVKDAVYTVV
543	1893		4152	(30)	l.,	SSH
343	1893	Α	4153	678	11	TISYPQCLTQMYFLISFANVDTFLLPIMALDH
						YVAICSALQ*CSITTP/ELCQGLPVLA*AGSSLIS
		ł				PVHTVIMSRLAFCSSAQISHFYRDAYLLMKIA
					İ	CSHT*\NQHVFLGAVVLFLAPCALILVSYIRIA AAILRIPSPTRRRKACSICSSHLSLVTLFYGTV
	1					LGICI*PPDSFSAQDAIATIMYTVVTSMLNPFIY
						SLMNKEVQEAVRRLFSRGSHSSWCW
544	1894	A	4158	3	538	LLYAQAGVQ*LNLSSLQPQPAGLKQSSHPSLP
1						SSWDYRYSTPHPANFFVEMEFHHVAQAGLEL
						LGSGDLPTSTSHSAGITGV\SHHAPPRLISSEGS
<u> </u>				,		LLGHLLCLPMVFPLLCVFVLISSSLAGEEAAG
]						LRVQKLWPAVVLSHLPVCWFHCSGIWSEVIE
<u> </u>						LKVGREGHVLPWQAHVVEF
545	1895	Α	4160	1	412	HPLGLGLVPSEIFSPQDKKAADGSILAPARGE
						DLEAGLKGSFMDGRLQASVSVFRIQRVGSAM
						QDTASAMPCLPYYPTSHCFMAGGKSRSQGW
						ELELSGEPAPGWQVLAGYTYTQARYLRDASE ANVGQPLRPVDPR
546	1896	A	4174	1252	1190	FFQVFIFLFFKTEFHSCCPGAVQWHDLDSL
						QPPPPRFKGFSCLSLPSSWDYRHAPAHPANFV
				i		FLVETGFLHV\GQ\ASLELPTSGDTPAS\ASQSA
					 	GITGVSHHA*PRASGRRCW
547	1897	A	4176	3029	1	AGPDGLAAPASCQGARGQTRVPGAFSWLAP
						GSHHASEGLAPGVPPAGGVSAQELTAPPQEG
						WGLGAPPAAPRPESDEKRAGSDAVRSFSRGA
1	l	ł	ŀ	ł	}	RDSLGQRRLGGTRGAGPAGKGAQRTMGPAS
						GFHSFPPRPHQEPSPRSSCWQHLLWHCPWPQ
	i i					PSRLPRLTPAQLLQGPGVLAAPPGP*HVPGFL AQSPWPLPSGPRSP*DPLHQGALVPLPQGGSP
						HTAPHCLPSVLSPAIQQPLLPTAST/SSRSPPAS
1	Ì	ł	1	1	1	TMAPIPSALAVWEPAGSSPQLSSAPADSS/PLP
				1		ALPKVLPPWTQKPLLGCLCQSPLPLLSPPDOI/
				1	1	RCPPACSPAAASSFSFESQPCPSAPSKASPAPA
		ł	-	- I	[AL\IVGPHHPP*SQQPQSQSVHPHGPGGPQPPL
' l	ŀ	1	- 1	ł	ļ	AASSLFWMFCQPPPPHPQFLWHRPLPVTGKA
.	İ			l	1	LAS\PLCFRPAPGSLRQTPLPPQFHIPRPGLSAP/
	l				1	PPPASGTSDSSDSRSPSASAARVWPPA\SPPPP
				l	I	AARHRPHPPEYFLSPCPFSCGFPRLLGRPRRPQ
1	- 1	i	}	ŀ	j	ALQTPRAWDLPPGSSPAPLCSGPELP*APPPLP
1		- 1	ļ	ļ	İ	PFPRVA*LGSGHPPSAQVPGLW*RCV*GHPIP
	1	1		}	ŀ	RPVGHS*SGPPHSPPL*APPQAWPLELPPSRQC LQPLHLRAAQPLDPCCSLSPPGPPLPVPALPS
	1	l		ł		WPGRP*SPSPASSQPPYHAGLPGPQSSPLPPGL
1	1	ļ	ļ	l	ļ	POLPSLRSGSOOPLLFFOCPGPGAVWGKGSPO
			1			PLSPHPPPP/ARTOTFPVASRSLSPGTAPYSVCL
1	1		1	ł	1	TPSRSASSLPEVVLASSLPKIPQSSGS\PLGPTSP
l	Ì	i	-	l	Ì	MP*CFHRPSPPLP/LSSPFPA\LRPQAPQFPLHLP
1						
- 1	- 1	- 1	1	ł		P*PPAPSPGCPLPPLAQQHQPSPPSPHARSTLT PPLWPSLALLP*PLPPPPPVPSFSASLLCSLPAH

SEQ ID	SEQ ID	1 Man	Lero	1 8 22 4 3	18 9.11	
NO: of	NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1 1100	in in	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	i	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
ł	ļ	1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ł		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	İ		1	peptide	'	/=possible nucleotide deletion, \=possible
				sequence	}	nucleotide insertion
						GTPASPGLGRSCLGKPQTLPWISFWPPSGRLA
						PGTWQPW/PVSPAPLSCLSAWDPWELPSPOPO
		1]	İ	VCSTAELPTSCLLSSPGP\PAFQPPRFGCL*GPP
	•			1		GPPGLPPLQSSLSFPPPPPPVPQPPAPPALQWG
510	1000	 	4:00	10000	<u> </u>	LHLPGGRTK
548	1898	A	4180	2369	844	RIHREEDFQFILKGIARLLSNPLLQTYLPNSTK
	İ	ļ	İ		j	KIQFHQELLVLFWKLCDFNKVGQPRGALQGD
				ł		GEQLPQ*PGGRDSVRLRGVGQSCPSLELSPLG
				Í		PSPHP*KFLFFVLKSSDVLDILVPILFFLNDAR
		[ĺ	ADOSRVGLMHIGVFILLLLSGECNFGVRLNKP
	İ					YSIRVPMDIPVFTGTHADLLIV\VFHKIITSGHQ
		ļ				RLQPLFDCLLTIVVNVSPYLKSLSMVTANKLL HLLEAFSTTWFLFSAAQNHHLVFFLLEVFNNI
		l	1	ŀ	ļ.	IQYQFDGNSNLVYAIIRKRSIFHQLANLPTDPP
•		l	1 1		l	TIHKALQRRRRTPEPLSRTGSQGGAPPWRAPA
		l				PLPLQSQAPSRPVWWLLQALTS*PRSPRCQR
						MAPCGPWNLSPSRAWRMAARLRGSPARHGG
						SSGDRP/HSSASGQWSPTPEWVLSWKSKLPLQ
		i	1			TIMRLLQVLVPQVEKICIDKGLTDESEILRFLQ
			1			HGTLVGLLPVPHPILIRKYQANSGTAMWFRT
						YMWGVIYLRNVDPPVWYDTDVKLFEIQRV
549	1899	A	4191	858	321	LPWQRLGVLLSRGKMAVTGWLESLRTAQKT
			1 1			ALLQDGRRKVHYLFPDGKEMAEEYDEKTSE
				,		LLVRKWRVKSALGAMGQWQLEVGDPAPLG
						AGNLGPELIKESNANPIFMRKDTKMSFQWRIR
						NLPYPKDVYSVSVDQKERCIIVRTTNKKYYK
550	1900	A	4192	1	1980	KFSIPDLDRHQLPLDDALLSFA\TPTAP IRHTGSDIAGVCGWLLLSGPCGVGLDLDSRLL
	1700		11,72	.	1700	GASAMRRSEVLAEESIVCLQKALNHLREIWE
						LIGIPEDQRLQRTEVVKKHIKELLDMMIAEEE
						SLKERLIKSISVCQKELNTLCSELHVEPFQEEG
]	ļ		ETTILQLEKDLRTQVELMRKQKKERKQE\LKL
						LQEQDQELC\EILCMPHYDIDSASVPSLEELNQ
						FRQHVTTLRETKASRREEF/VSSIKRQIILCME
	:		ļ l			ELDHTPDTSFERDVVCEDEDAFCLSLENIATL
l l]]			QKLLRQ\LEMQKSQNEAVCEG\LRTQI\RELW
			<u> </u>	1		DRLQIPEEEREAVATIMSGSKAKVRK\ALQ\LE
				1		VDRLEELEKCKTMKKVIEAIRVELVQYWDQC
	ļ		j			FYSQEQRQAFAPFCAEDYTESLLQLHDAEIVR
						LKNYYEVHKELFEGVQKWEETWRLFLEFER
ĺ	1			1	1	KASDPNRFTNRGGNLLKEEKQRAKLQKMLP
	İ					KLEEELKARIEL WEQEHSKAFMVNGQKFME
	1			_]	_ [YVAEQWEMHRLEKERAKQERQLKNKKQTET EMLYGSAPRTPSKRRGLAPNTPGKARKLNTT
I	- 1			-1	· · ·]	TMSNATANSSIRPIFGGTVYHSPVSRLPPSGSK
1	ĺ			i	1	PVAASTCSGKKTPRTGRHGANKENLELNGSI
	l			1	İ	LSGGYPGSAPLQRNFSINSVASTYSEFADPSLS
1	_			1	Í	DSSTVGLQRELSKASKSDATSGILNSTNIQS
551	1901	À	4194	3	1008	AWHEGLVSSPAIGAYLSASYGDSLVVLVATV
- 1	Í	1	1	i	1	VALLDICFILVAVPESLPEKMRPVSWGAOISW
					i	KQADPFASLKKVGKDSTVLLVCITVCLSYLPE
		ļ		l	ŀ	AG\QYSSFF\LYLR\QVIGFG\TVKIAAFIAMVGI
ļ		- 1		ļ		LSIVAQTAFLSILMRSLGNKNTVLLGLGFQML
ł	i	ł		}		QLAWYGFGSQAWMMWAAGTVAAMSSITFP
				İ	1	AISALVSRNAESDQQGVAQGIITGIRGLCNGL
- 1				İ	ļ	GPALYGFIFYMFHVELTELGPKLNSNNVPLQ
l l	1	1			1	GAMBORDEL ECACIM MOELVAL PIDEVOVAC
	I	į			i	GAVIPGPPFLFGACIVLMSFLVALFIPEYSKAS
-]			j	GVQKHSNSSSGSLTNTPERGSDEDIEPLLQDS SIWELSSFEEPGNQCTEL

No. of peptide coide of U.S.S.N operation of the company of the control of U.S.S.N operation of the company of the control of U.S.S.N operation of the control of the contr	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence oide seq uence very sequence sequence sequence sequence sequence very sequence v					1		Annie acid sequence (A=Alanine C=Cysteine,
Sequence	nucl-	peptide					F=Phenylalanine G=Glycine H=Uictidine
uence 914	cotide	seq-		USSN			I=Isoleucine K=I vsine I =I eucine
1914 majino first amino acid residue of peptide pequence 1910 1	seq-	испсс	i	09/496	correspondi		M=Methionine N=Asparagine P=Proline
amino acid residue of peptide sequence physical sequence of peptide sequence physical sequence of peptide	neuce		[914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine
residue of poption of poption (poption) (poption) residue of poption (poption) (poptio	ł		}		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan
Special Sequence				1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon
1902 A 4197 2 14302 APPPAPAGRONONAAAAAELRGAR EPAPARRAGTMADGGEGEDEIQPLXTIDEVV LQCTATIKEGQKICLAAEGGARNICLESTS NSKNVPPDLSICTFVLEQSLSVRALQEM_ANT VEKSEQVOVPEKWERMKTAQGGERITIL VGHABLRHSVSGWYLCCLSTSRSSTDKLAED VGLQEDTTGAECWWIHPASKQNSIGEGWRY VGEDLIL VSVSSSRYJLHSVGNGSLHVDAAF QGTUWSVAPISSGSSLAAGVIJGGDVRLLH GIMDECTTVPSGEIGEGRRYTVIYEGGAVS VGABLLWASPISSGSSLAAGVIJGGDVRLLH GIMDECTTVPSGEIGEGRRYTVIYEGGAVS VGABLLWASPISSGSSLAAGVIJGGDVRLLH GIMDECTTVPSGEIGEGRRYTVIYEGGAVS VGABLLWASPISSGSLAAGVIJGGDVRLSTA FITRSSKERLDVGVAREUDGMGTSEKY GDS VCYQIQVIDGUMTYGSAVIPSGSWRSGNG KAMHEGGIMDDGISLSRSQHEESTARVIRS TVFLFARFREGDALSKKAKASTVD_IPSVSL SLQDLIGYFIPPDEHLEHEDKONRLRALKNF QMFQEEGMINULECIDIALHVYSSAAHRAD VAGREAGESWKSILNSYSLLLAAIRGNRKN CAQFSGSLDWUISRLERLEASSGLEVLUCV VESPEALMIREGIRKSISLDKHGRNIKVLD VLCSLCVCHGVAVRSQHCICONLLPGRAC VESPEALMIREGIRKSISLDKHGRNIKVLD VLCSLCVCHGVAVRSQHCICONLLPGRAC VESPEALMIREGIRKSISLDKHGRNIKVLD VLCSLCVCHGVAVRSQHCICONLLPGRAC VESPEALMIREGIRKSISLDKHGRNIKVLD VLCSLCVCHGVAVRSQHCICONLLPGRAC VESPEALMIREGIRKSISLDKHGRNIKVLD VLCSLCVCHGVAVRSQHCICONLLPGRAC VESPEALMIREGIRKSISLDKHGRNIKVLD VLCSLCVCHGVAVRSQHCICONLPGRAC VESPEALMIREGIRKSISLDKHGRNIKVLD VLCSLCVCHGVAVRSQHCICONLPGRAC VESPEALMIREGIRKSISLDKHGRNIKVLD VLCSLCVCHGVAVRSQHCICONLPGRAC VESPEALMIREGIRKSISLDKHGRNIKVLD VLCSLCVCHGVAVRSQHCICONLPGRAC VESPEALMIREGIRKSISCLDKHGRNIKVLD VLCSLCVCHGVAVRSQHCICONLPGRAC VESPEALMIREGIRKSISCLDKHGRNIKVLD VLCSLCVCHGVAVRSQHCICONLPGRAC VESPEALMIREGIRKSISCLDKHGRNIKVLD VLCSLCVCHGVAVRSQHCICONLPGRAC VESPEALMIREGIRKSISCLDKHGRNIKVLD VLCSLCVCHGVAVRSQHCICONLPGRAC VESPEALMIREGIRKSISCLDGRAC VESPEALMIREGIRKSISCLDSEALAFTOPTUT GUQDVKRRTPETTVT GUQDVKRRTPETTVT GUGDVKRRTPETTVT GUGDVKRRTPETTVT GUGDVKRRTPETTVT GUGDVKRRTPETTVT GUGDVKRRTPETTVT GUGDVKRRTPETTVT GUGDVKRRTPETTVT GUGDVKRRTPETTVT GUGDVKRRTPETTVT GUGDVKRRTPETTVT GUGDVKRRTPETTVA GUGDVKRRTPETTVA GUGDVKRRTPETTVA GUGDVKRRTPETTVA GUGDVKRRTPETTVA GUGDVKRRTPETTVA GUGDVKRRTPETTVA GUGDVKRRT	}		ļ		peptide	_	/=possible nucleotide deletion. \=possible
EPAPARROTMADGOEGEDEIGETDDEVV LQCTATHIREQOKLCLAAEGFORICCTLESTS NERNYPPLSICITYLEQSISVRALQEMI AMT VEKSEGQVDVEKWKFMKKTAQGGGHRTIL YGHALLRHSYSGMYLCLISTRSSTRIKAFD VGLQEDTTGEACWWTHPASKQRSEGEKVR VGDDLLVSVSSERFYLHLSYGNGSLHVDAAF QQTLWSVAPISSGSEAAQGYLIGGDVLRLLH GIMDECLTVPSGEHGEGRRTHYBGGAVS VHARSLWRLETLRVAWSGSHRWGOPFELR HYTTGKYLSLIMEDKALLIMDEKEADVKSTA FIFRSSKEKLDVGVRKEVDOMGTSEIKYGDS VCYIQHVDTGLWLTYGSVDYKSVMGSIGR KAMMHEGHADDGISLSRSQHEESRTARVIRS TYFLFRFIRGLDALSKKAKASTVDLPIESVSSI SLQDLGYFHPPOEHLEIEDXQNRLRAKKNO QNLPQEEGMIN.VLECIDRLHVYSSAAHFAD VAGREAGESWKSINSSISVLELAALIRGRKN CAQFSGSLDWLISRLERLASSGILEVLHCVL VISPEALNIKEGHIKSSILLDKHGRRHKWLD VLCSLCVCHGVAVRSNQHLLONLLPGRDLL LQTRLVSHYSSMRPHIFLGVSGSGAYKKWY YELMYDHTEFFYTAEATHLRVGWASTEGYSP YFGGGEEGWGGNGYGDDLFSYGFGGILUNGG CLARTVSSPNQHLLRTDDVISCCLDLSAPSISF RNOQPYQAMEENNIDGLFFVVSSAGIKV RRLLGGRHGEFFLPPPGVAPCYEAVLFKERL KVEHSREYYGGETVTRDLGFPVSFAGIKV RRLLGGRHGEFFLPPGVAPCYEAVLFKERL KVEHSREYYGGETVTRDLGFPVSFAGIKV RRLLGGRHGEFFLPPGVAPCYEAVLFKERL KVEHSREYYGGETVTRDLGFPVSFAGIKV RRLLGGRHGEFFLPPGVAPCYEAVLFKERL KVEHSREYYGGETVTRDLGFPVSFAGIKV RRLLGGRHGEFFLPPGVAPCYEAVLFKERL VEKKMERYYGETVTRDLGFPVSFAGIKV RRLLGGRHGEFFLPPGVAPCYEAVLFKERL VEKKMERYYGETTTRDLGFPVSFAGIKV RRLLGGRHGEFFLPPGVAPCYEAVLFKERL VEKKMERYYGETTTRDLGFPVSFAGIKV RRLLGGRHGEFFLPPGVAPCYEAVLFKERL VEKKMERYYGETTTRDLGFPVSFAGIKV RRLLGGRHGEFFLPPGVAPCYEAVLFKERL VEKKMERYYGETTTRDLGFPVSFAGIKV RRLLGGRHGEFFLPPGVAPCYEAVLFKERL VEKKMERYYGETTDLGFFVSFAGIKV RRLLGGRHGEFFFLPPGVAPCYEAVLFKERL VEKKMERYYGETTTRDLGFFVSFAGIKV RRLLGGRHGEFFFLPPGVAPCYEAVLFKERL GUGDWYGNENGOPDCLGGBERAFECG ERNYNLOMSLETIKTTLALGCGVGISDEHAE DEVKKMERYGETTGGSVFAPMOLISTRICT GLOEGVERFRAPTORITOTTWLSKRILFTGC GEFFFRATARRYRLYFTTLDDRTKKSNKDS LRAVYRTLGGYGFPADGGRRNNGGFTC GFFPCSLGVAQVGRNNFGKDVSTLKFFTTC GLOEGVERFRATNTROTTMINLSKRILFFTTC GLOEGVERFRATNTROTTMINLSKRILFFTTC GLOEGVERFRATNTROTTMINLSKRILFFTTC GLOEGVERFRATNTROTTMINLSKRILFFTTC GLOEGVERFRATNTROTTMINLSKRILFFTTC GLOEGVERFRATNTROTTMINLSTRALFOPLOV PSNHEHIEVTRUGGTDDSSPCLKYTOTTGDEKGKVHE SIKRSNCTVVGCGGRRNNGLEGG VVDAASGLLTFTANGKELSTYVQCTGSRILFQE RSTAVSTSISHECYGETYPOTTMINL					sequence		nucleotide insertion
EPAPARRGTMADGGGEDEJGFLRTDEVY LQCTATHIKEQQKLCALAGEGGNICTLESTS NSKNVPPDLSICTFVLEQSLSVRALQEMLANT VEKSEQVDVEK WEKFMKKTAQGGGRITTLL YGHALLIKHSYGMYLCCLSTSRSSTDKLAFD VGLQEDTTGBAC WYTHHASKQUSSGEKVR VGDDLLLVSVSSERYLHLSYGNGGLHVDAAF QQTLWSVAPISSGSEAAQGYLIGGDVLRLLH GHMDECLTVPSGGHEGERRTVHYEGGAVS VHARSLWALETLRVAWSGSHRWGOFFRLR HYTTGKYLSLEMEDKNLLLMDREEADVKSTA FFRESKEKLDVGVRKEVDGMGTSEIKYGDS VYQHVDTGLWLTVGSVDVSSYMGSIQR KAMMHEGHMDDGISLSRSQHEESRTARVIRS SIQDLGYFHPPDEHLEHEDKONRLRALKNR QNLFOGEGMIN, VLECDERLHVYSSAAHFAD VAGREAGESWKSILNSLYELLAALRGNRKN QNLFOGEGMIN, VLECDERLHVYSSAAHFAD VAGREAGESWKSILNSLYELLAALRGNRKN CAGPSGSDWLISRLERLEASSGLEVHLYVL VESPEALNIKGGHKSIISLLDKHGRNRKVLD VLCSLCYCHOVAVRSOPQLITCDILLPGODLL LQTRLVSHVSSMRPNIFLGVSGGSAQYKKWY YELMYDHTEFVTAEATHLRVGWASTEGVSP YFGGGEWGGNGGYGDLIFSYGFGGLH.WSG CLARTVSSPQHLLRTDDVISCCLDLSAFSISF RNGQPVQGMEENNIDGLFFPVYFSAGKW RFLLGGRHGEFKFLPPFGYAFCYEAVLFKEKL KVEHSBERYXGERTYTRDLGFTVSTGAGKF RPIPVTSGIVLPPHLERREKLAENIELLWMN KIELGWGYGFWEDNIKRGHFLVFSSLEGE ERNYNLQMSLETLKTLLALGRIELLWMN KIELGWGYGFWEDNIKRGHFLCVFSSLEGE ERNYNLQMSLETLKTLLALGRIELLWMN KIELGWGYGFWEDNIKRGHFLVFSSLEGE ERNYNLQMSLETLKTLLALGCFWGSDEHAE DVKKKMKLPRNYQLTSGYKFAPMLSFEILT PSQEAMYDKLAENAHINWARRIRGGWTY GIQDVKNRNPRILVTYTDLDGRTXSSTNSDS LREAVRTLLGGYGYFFDGGLMAGEFELT PSQEAMYDKLAENAHINWARRIRGGWTY GIQDVKNRNPRILVTYTDLDGRTXSSTNSDS LREAVRTLLGGYGYGFWEDNIKSNSDS LREAVRTLLGGYGYGRAPDHAAAEVCS GTGFRFREFAETTYAYACRGWYFEFETVTA GGMRVGWRRGPGCYPDQELGSDERAFAFDGF KAQRWHQGNNFRKNDVAYACRGWYFEFETVTA GGMRVGWRRGPGCYPDQELGSDERAFAFDGF KAQRWHQGNNFRKNDVSTLKYFTIC GLOEGVEFFANNINDTIMWLSKRLPGFLQV PSNHEHIEVTRIDGTIDSSPCLKYTYGSFGSQN NSTIDMFYRLSMPEGAEVSKTVAGGLPGAG LFOPKNDLEDYDADSDFEVLMKTAHGGILVP DFMKDLEATYFTTHRNIKYYACRGWYGWTT DFHQYTTYTYSVRIFTGGPANVWGWTT DFHQYTTYTYSTSISIARLTEDVLADDRDDY DFMKGTTYTYSVRIFTGGPANVWGWTT DFHQYTTYTYSTSISIARLTEDVLADDRDDY DFMKGTTYTYSVRIFTGGEPANVWGWTT DFHQYTTYTYSTSISIARLTEDVLADDRDDY DFMKGTTYTYSVRIFTGGEPANVWGWTT DFHQYTTYTYSTSISIARLTEDVLADDRDDY DFMKTTHAMFELLTGELLGHERKYHWFHT EFKSTILTPDOKKHCIJOFISISTSRRWQFH KVDAASGLLTTAARLMMNEVPLYMT EFKSTILTPDOKKHCIJOFISISTSRRWGHE SKSFNYSISN	552	1902	A	4197	2	14302	ARPPPAPGSROOKOKAAPGAAAAELRGAR
LQCTATHEEQQKLCLAAEGGGRETLL NEKNYPPDLSICTFYLGOSLYNKALQEGGHRTLL YGHALLKHYSGMYLCLISTRSSTENLAFD VGLQEDTTGEACWNTHFASKQRSIGGEKYR VGDDLLVSVSSERYLHSYGNGSLHVDAAF QOTLWSVAPISSGSEAAQGYLIGGDVLRLLH GHMDECLTVPSGEHGEBORT YHYPTGAAYS VHARSLWRLETLRVA WSOSHRWGOPFRLR HYTTCKYLSLMEDKNILLMKREFADVKSTA FFFRSSKERLDVGYRKEVDGMGTSEIKYGDS VGYQHVDTOLWLTYGSVDVKSVRMGSIQR RAMMHEGHNDDGISLRSKHEESKTARVIRS TYFLFNETRGLDALSKKAKASTYDLPIESVSL SLQDLIGVFHPPDEHLEHEDKONRLFALKINR QMLFQEEGMMLVLECIDRLHVYSSAAHFAD WGREAGESWKSLINSIS, YELLAALIRGRIKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALMIKEGHIKSISLLDHGRIFHKVLD VLCSLCVCHGVAVRSNOPHLICDNILLFORDLL LQTRLVSHYSSMRPHIFLGVSEGSAGYKKWY YELMYDHTEFFYTABATHLRVGWASTEGYSP YNGGGEWGGNGVGDDLFSYGFGGILUNGG CARTYSSPNOHLLRTDDVISCCLDLSAPSISF RNGQPVGMFENNDDLFFPVTSSGIKV RFLLGGRHGEFKFLPPPGVAPCVEAVLPKERL KVEHSREYYAGERLYTDLLGPTVELTOAAFT PIPVDTSGIVLPPHERIRELLABHELWVMN KIELGWYGFVEDDNKRGHFCLVEFSKLFEG ERNYNLQMSLETILKTLLLAGCHVGISDEHAB DKVKKMKLPRNYQLTSGYKRAPMDLSFRLT PSGEAMYDLLABAHRNWARDRIRGGWTY GIQQDVKNRRNFRLYPTTELDDRTKKSNKDS LRAVKTILLOGYGNLEAPDGHARAREVCS GTGERFRIFABEKTY AVKAGRWYFEFETYTA GDMRWGWSRFGCOPDGLGSDERAFBOG KAQRWHGNEHYGRSWQAGDVVGCMYDM NEHTMAFTLLGGTHDSSPLLKYTTCL GDMRWGWSRFGCOPDGLGSDERAFBOG KAQRWHGNEHYGRSWQAGDVVGCMYDM NEHTMAFTLLGGTHDSSPLLKYTTCL GLGGGYFPANTINTDTIMWLSKLFOFLOY GPPVCSLGVAQVGRNRHGKDVSTLKYFTC GLGGGYFPANTINTDTIMWLSKLFOFLOY PSNHEHIEVTRUGGTDSSPLLKYTOKSFGSQN SINDMFYRLSPFEODPLARAFBOCH KAQRWHGNEHYGRSWQAGDVVGCMYDM NEHTMAFTLLOGGTHUSSSELLKFOFLOY PSNHEHIEVTRUGGTDSSPLLKYTOKSFGSQN SINDMFYRLSPFECHERICH AFDGLOPAG LFGPKNDLEDVAASGLPAG LFGPKNDLEDVAASGLPAG LFGPKNDLEDVAASGLPAG LFGPKNDLEDVAASGLPAG LFGPKNDLEDVARSHWAGLEGGLSTSLYPHOP PROVESCHARTHPOORMARARVCS GHERSTYYSTYSSISARLTEDVLADDRDDY DFHMYTTYYSYLEDFORLESTYTYVPTESTKLPF AVFAQATSPNYOFLEGRIKYNWEPSERLO RYDLESTRYSTYMPOFLEGRIKYNWEPSERLO RYDLESTRYSTYMPOFLEGRIKYNWEPSERLO RYDLESTRYSTYMPOFLEGRIKYNWEPSERLO RYDLESTRYSTYNOGLISTSLRFRMOP RYDLESTRICETERSELLAGHERICHOLISTSLRFRMOP RYDLESTRICETERSELLAGHERICHOLISTSLRFRMOP SSSFYSISINSECYYSPFFULDLKSTIONLE]						EPAPARRRGTMADGGEGEDEIOFLRTDDEVV
NSKNYPPLSICITYLEQSLSYRALQEMIANT VEKSEGQVDVEKWEYMMKTAQGGIRITLL YGHALLRHSYSGMYLCCISTSRSSTDKLAED VGLQEDTTGBACWWITHPASKQRSGEKVR VGDDLILVSVSSERYLHLSYGNGGLHVDAAF QOTLWSVAPISSGEAAQGYLIGDVAULH GIMDECLTYPSGEHGEQRRTVHYEOGAVS VHARSLWRLETLRVAWSGSHRWGQFFULR HYTTGKYLSLMEDKNILLMKKEREADVKSTA FFFRSSKENLLLMKOKEADVKSTA FFFRSSKENLLLMKOKEADVKSTA FFFRSSKENLLLMKOKEADVKSTA FFFRSSKENLLLMKOKEADVKSTA FFFRSSKENLLLMKOKEADVKSTA FFFRSSKENLLLMKOKEADVKSTA FFFRSSKENLLLMKOKEADVKSTA FFFRSSKENLLLMKOKEADVKSTA FFFRSSKENLLLMKOKEADVKSTA FFFRSSKENLLLMKOKEADVKSTA FFFRSSKENLLLMKOKEADVKSTA FFFRSSKENLLLMKOKEADVKSTA FFFRSSKENLLLMKOKEANSON KAAMHEGHMODGISLSRSQHEESRTARVIRS SLQDLGYFHPPDEHLEHEDKONRLRALKNR QNLFOGEGMINLVLECDEURLHYVSSAAHFAD VAGREAGESWKSILNSLYELLAALIRGNRKN QNLFOGEGMINLVLECDEURLHYVSSAAHFAD VAGREAGESWKSILNSLYELLAALIRGNRKN CAOFSGSLOWLISBLERLEASSGIE US-HCH-VL VESPEALNIKEGHIKSIISLLDKHGRNRKVLD VCSLCYCHOVAWRSOMPLLOPEDLPHOPDL LQTELVHHVSSARPHIFLGVSGSAQYKKWY YELMVDHTEPFVAAEATHLRVGWASTEOSY PYOGGEEWGONGVGDLISYGFGGLHLWSG CLARTVSSPQHALLRTDDVISCCLDLSAPSISF RNGQPVQOMFENNIDGLIFPDVYSFSGIKV RRLLGGRHGERKILPPPGVARCYEAVLPKEKL KVEHSREYYQERLYTRDLGPTVSLTOAAFT PIPVDTSQIVLPPHLERIRELLAENHELWYMN KIELGWGYGPVBONKRQHPCLVFSSLYEQ ERNYNLQMSLETLKTTLLAIGCHVGISDEHAE DEVKKMKLPRNYQLTSGYKPAMDLSFELLT PSQEAMVDKLAENABNWARDRIRGGWTY GIOQDVKNRNPRILVTYTDLDGRTKSSINED LREAVRTLLGYGYNLEAPDODHAARAEVCS GTGFRERIFFAEEKTVAYACAGWYFEFTIVTA GOMRYGWSRGCOPPOQELGSDERAFAFDGF KAQRWDKLAENABNWARDRIRGGWTY GOMRYGWSRGCOPPOQELGSDERAFAFDGF KAQRWDKLAENABNINWARDRIRGGWTY GOMRYGWSRGCOPPOQELGSDERAFAFDGF KAQRWDKLAENABNINWARDRIRGGWTY DEMONSTYTYSVERIFFTOTT GLORGVEFFANTINTOHINKNIKKLPFTIC GLOEGVEFFANTINTOHINKNIKKLPFTIC GLOEGVEFFANTINTOHINKNIKAGENSILD SKRSNCYAVCAGESEAFVSKTVAGGLOPGG LFOPKNIEHEVTRUDGTIDSSPCLKYTGKSFGKON NIDIMFYRLSMPECAEVVSKTVAGGLOPGG LFOPKNIEHEVTRUDGTIDSSPCLKYTGKSFGKAD SKRSNCYAVCAGESASPCGORNNOHGUEG VVDAASGLLTHAMCKELSTVYOVPESTKLP DFHQYDTGFDLDRWTTYTYTI.GDEKGKVHE SKRSNCYAVCAGESASPCGORNNOHGUEG VVDAASGLLTHAMCKELSTVYOVPESTKLP AVFAQATSTAYSTONGELEGE SKRSNCYANDEADSTRAENABMNOHEN RYMLBLITGREENAKYMPGLIR AGYYDLLDHLSSYATARAMMAMSTU	1 1						LQCTATIHKEQOKLCLAAEGFGNRLCFLESTS
VEKSEGOVDVEKWKFMMKTAQGGGHRTLL YGHALLRHSYSGMYLCLISTRSSTRIKLAFD VGLQEDTTGEACWNTHPASKQRSEGEKVR VGDDLLVSVSSERYLHLSYMGSLHVDAAF QQTLWSVAPISSGSEAAQGYLIGGDVLRLIH GHMDECLTYPSGEGHEGORRTYHYEQOAVS VHARSLWRLETLRVAWSGSHIRWGOFFRLR HYTTGKYLSLMEDKNLLLMDKERADVKSTA FFRSSKEKLDVGWKSKVDGMGTSEIKYGDS VCYQHVDTGLWLTYQSVDVKSVRMGSIQR KAMMHRIGHMDDGISLRSQHEESFRAVIRS TYFLFNRFIRGLDALSKKAKASTVDLPIESVSI SLQDLIGYPHPDEHLHEIDKONRLBRALKIN QNLFQEEGMINLVLECIDRLHVYSSAAHFAD VAGREAGESWKSLINSLYSLLAALIRGRIKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPBALMIKEGHKSISLLDKHGNRHKVLD VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY YSLMVDHTEFFVTABATHRLVGWASTEGYSP YPGGGEEWGGNGYGDDLFSYGFDGLHLWSG CARTVSSPNQHLLETDLDVSCCLDLSAPSISF RNQQPVQGMFENFNIDGLFFPVSFSAGIKV RFLLGGGHGEKFLFPPVSFSAGIKV RFLLGGGHGEKFLFPPVFSFSAGIKV RFLLGGGHGFKFLPPFVAFCRAVLFKEL KVEHSRSYKOERTYTRDLLGPTVSLTQAAFT PIPVDTSQVLPPHLERREKLAENDELLWYNN KIELGWQYGPVKDDMKRQHPCLVFERSLDPO ERNYNLOMSLETLALIGCHVGISDEHAE DKVKKMKLPRNYQLTSGYRPAMDLSFIRLT PSQEAMVDKLAENAHNWARDRIRGGWTY GIQDVKNRRNFRLVYTTLDDLGKCHSENFLT PSQEAMVDKLAENAHNWARDRIRGGWTY GIQDVKNRRNFRLVYTTLDDLGKGNJEFRLT PSQEAMVDKLAENAHNWARDRIRGGWTY GIQDVKNRRNFRLVYTTLDDLGKGNJEFRLT PSQEAMVDKLAENAHNWARDRIRGGWTY GIQDVKNRRNFRLVYTTLDDLGKGNJEFRLT PSQEAMVDKLAENAHNWARDRIRGGWTY GIQDVKNRRNFRLVAYAGAGWYFFETVTA GDMRVGWSRFGCQPDQELGSDERAFATDGF KARWHQCNEHYGRSWQAGDVCGMVDM NEHTMAFTLNGEILLDDSGSLAFKDPDVGD SMRVGWSRFGCQPDQELGSDERAFATDGF KARWHQCNEHYGRSWQAGDVCGMVDM NEHTMAFTLNGEILLDDSGSLAFKDPDVGD SMRVGWSRFGCQPDQELGSDERAFATDGF KARWHQCNEHYGRSWQAGDVCGMVDM NEHTMAFTLNGEILLDDSGSLAFKDPDVGD SMRTURSTYYSVSUFFRQEANNYGRNGENSCHAYPTIC GLQBGYEFFANTNEDTIMVLSKLEPQFLQV PSNHEEIBVYTRDDGTDDSPCLKVYQGSGSQN SMTDMFYRLSMFECAEVSKTVAGGLPGG KARWHGNEHYGRGPANTWGWGTTS DFHQYTTGFDLDRYRTYTVTLGDEKGKVHE SKRSNCYWACAGSSNYGGRGCQCMYDM KEHTMAFTLNGEILLDDSGSLAFKDPDFD DFMCTSTYYSVSUFFRQEPANVWGWTTS DFHQYTTGFDLDRYRTYTVTLGDEKGKHEIL SKRSNCYWACAGGSSROPGCGNNGGLIGC VVDAASGLLTFIANKYMPGLER RNDULESTGQERLKYMFVPMT EETKSITLFPDDNKKHGLPGGLSTSLRPMOTP							NSKNVPPDLSICTFVLEOSLSVRALOEMLANT
YGHALLAHSYSGMYLCCLSTSRSSTDKLAD VGLQEDTTGEACWTHPASGQRSGESEVR VGDDLLVSVSSERYLHLSYGNGSLHVDAAF QOTLWSVAPISGSEAGQVYLGGDVLRLH GHMDECLTVPSGEHGEEQRRTVHYSGGAAF QOTLWSVAPISGSEAGQVYLGGDVLRLH GHMDECLTVPSGEHGEEQRRTVHYSGGAAF WHARSLWRLETLRVAWSGSHRWGQPFRLR HYTTGKYLSLMEDKNLLLMMEKADVKSTA FFRSSKELDVGVRCEVUGMGTSEIKYGGDS RAMMHEGHMDDGISLSRSQHEESRTARVISGS KAMMHEGHMDDGISLSRSQHEESRTARVISGS KAMMHEGHMDDGISLSRSQHEESRTARVISGS KAMMHEGHMDDGISLSRSQHEESRTARVISGS TVFLIFNRFIRGLDALSKKAASTVDLPIESVSL SLQDLIGYTHPPDEHLEHEDKQNRLRALKN QNLFQEEGMINLVLECIDRLHVYSSAAHFAD VAGREAGESWKSILNSLYELLAALIRGNRKN QNLFQEEGMINLVLECIDRLHVYSSAAHFAD VAGREAGESWKSILNSLYELLAALIRGNRKN CAQFSGSLDWLISRLERLEASSGILDHGRNHKVYL VLCSLCVCHGVAVRSNOHLICDNLLPGRDLL LQTRLVNHVSSMRPNIFLGVSEGSAQYKWY YELMVDTTEPPVTAEATHLRVGWASTEGYSP YPGGGEEWGGNOVGDDLFSYGFDOLH WSG CLARTYSSPNQHLLRTDDVISCCLDLSAPSISS RINGQPVQMEENFINDLIFPTVSFSAGIKV RRLLGGRRIGEFKLPFPGYAPCYEAVLPKKL KVEHSFEYKGERTYTDALLGFTVSFSAGIKV RRLLGGRRIGEFKLPFPGYAPCYEAVLPKKL KVEHSFEYKGERTYTDALLGFTVSTOAAFT PIPVDTSQIVLPPHLERIREKLAENHELWWM KIELGWYGGVFUDNKRQHPCLYSTAPADLSFRILL PSQEAMVSLAENHANDARDRGQWTY GIQQDVKNRRNPRLYPTTDLDETSKLAENHELWWM KIELGWYGGVFUDNKRQHPCLYSTKAPADLSFRILL PSQEAMVSLAENHANVARDRRGGWTY GIQQDVKNRRNPRLYPTTDLDRFSKLPG ERNYNLQNSLETLKTLLALGCHVGISDEHAL DEVKKKMLEPKNVQLTSGYKPAPADLSFRILL PSQEAMVSLAENHANVARDRRGGWTY GIQQDVKNRRNPRLYPTTDLDRFSKLPG ERNYNLQNSLETLKTLLALGCHVGISDEHAL PSQEAMVSLAENHANWARDRRGGWTY GIQQDVKNRRNPRLYPTTDLDRFSKLPG GDMRVGWSRPGCQPDQELGSDERAPAFDGF KAQRWYGGNEHVFRSWQAGDVVGGMVBM NEHTMMFTLNGEILLDSGSLAFAFDGF KAQRWYGGNEHVFRSWQAGDVVGGMVBM NEHTMFTLNGEILLDSGSLAFAFDGF KAQRWYGGNEHVFRNHDLYAGGLPGGAG SINDIDMFYRLSMPFLAFFNNHDLYAGGLPGGAG SINDIDMFYRLSMPFLAFFNNHDLYAGGLPGGAG HPVCSLGVAQVGRMPGEGNSTLKYTTICL GLQEGFSPFAVRINGLIFMCAGEPSRLV PDRVOKKEATKEFTNNHDLYAGGLPGGAG SINDIDMFYRLSMPFLAFFNNHDLYAGGLPGGAG SINDIDMFYRLSMPFLAFFNNHDLYAGGLPGGAGA SINDIDMFYRLSMPTGCAFFSKTVAGGLPGGAG SINDIDMFYRLSMPTGLFSTNNGLIFMC RFLLRTKFDYTSTBGGTIDSSCLKVTQKSFGGGAG SINDIDMFYRLSMPTGLGFSRNFQGCFTNNGLIFMC WYDAASGLLTFANGKLESTYYQVEPSTKLF AVFAQATSPNYOGFELGRIKNTWMFLSAGLFEA KYDVILETTGGELKFTHYTHLXYAVCAL							VEKSEGQVDVEKWKFMMKTAOGGGHRTLL
VGLQEDTTGEACWWTIIHPASKQRSSGERAV VGDDLLLYSVSSRYHLBLYSVGNSLHVDAAF QQTLWSVAPISSGSEAAQOYLIGGDVIRLIH GHMDECLTYPSGEHGEORRYTHYSGQAVS VHARSLWILETLRVAWSGSHRWGQPFRLR HYTTGKYLSLMEDENLLLMDKEKADVKSTA FIFRSSKEKLDVOKKEVDJGMGTSEIKYGDS VCVIGHVDTGLWLTYQSVDKSVRMGSIGG RAMHEGHMDDGSLSSRSQHEESRTARVIRS TVFLFNRFIRGLDALSKKAKASTVDLPIESVS SQDLIGYFHPPGDEILEHEKOKNRRALKNR QNLFQEEGMINLVLECIDRLHVYSSAAHFAD VAGREAGESWKSLNSLYELLARGRKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALMIREGHIKSIISLLDKHGRYNKVL VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL LQTRLVAHVASSMRPNIFLGVSEGSAQYKKWY YELMVDHTEPFVAFAATHLRVGWASTEGYSP YYGGGEEWGGNGVGDDLFSVGFDGLHLWSG CLARTVSSRNQHLICTDDVISCLDLSPSISS RNQQPVQGMFENFINDGLFFFVVSFSAGKK RFLLGGRHGFEKFPPGVAPCFVALFKEKL KVEHSREYKQERTYTDLLGFTVSLTQAAFT PIPVDTSQIVLPHERIREKLAEHELVYMN KIELGWGYGPVRDDNKRGIPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAEBAHNVWARDRIRQGWTY GIQQDVRNRNFRIPLYPTLDTKKSNKDS LREANTILLGYGVNLEAPBOQHAARAEVCS GTGEFFRIFFAGKTYAVKAGRWYFEFETVTA GDMRYGWSRPGCQPDQELGSDERAFARDGG KAQRWHGGNEHVGRSWQGDVGGMVGDM NEHTMMFTLNGEILLDDSGELAFKDFDVGD GFPVCSLGVAQVGRMPGKDVGGMVGMDY SNIEBHEVTRIDGTMSSELAFKDFDVGD GFPVCSLGVAQVGRMPGKDVGGMVGMDY SNIEBHEVTRIDGTMSSELAFKDFDVGD FFNNELEHEVTRIDGTDSSPCLAVTOKGFRGND SNITDMFYRLSWPIECAEVFSKTVAGGLPGAG LFOPKNDLEDYDADSDFEVLMKTAHGHUP DRIVKDKAKATHENEVTRIDGTSSCLAVTOKGFRGND SNITDMFYRLSWPIECAEVFSKTVAGGLPGAG LFOPKNDLEDYDADSDFEVLMKTAHGHLVP DFNVGKAGFFRINGLEGC VVDAASGLITFIANGKELSTYYOVEPSTLLP DFHQYDTGFDLDRVRITVTVTLGEKKKVED SIKRSNCYMVCAGESMSFGGGRNNGLEIGC VVDAASGLITFIANGKELSTYYOVEPSTLLP AVFAQATSPNNYGGPCE AND NINGELFSK BHKNPYPCCPRLHVOFLSHVLWSRMMPNOFL KVDVSRISERGGWLVQCLDPLGFMSLHFEEN RKNPYPCCPRLHVOFLSHVLWSRMMPNOFL KVDVSRISERGGWLVQCLDPLGFMSLHFEEN RKNPYPCCPRLHVOFLSHVLWSRMMPNOFL KVDVSRISERGGWLVQCLDPLGFMSLHFEEN AHKNPYPCCPRLHVOFLSHVLWSRMMPNOFL KVDVSRISERGGWLVQCLDPLGFMSLHFEEN AHKNPYPCCPRLHVOFLSHVLWSRMMPNOFL KVDVSRISERGGVLVQCLDPLGFMSLHFEEN AHKNPYPCCPRLHVOFLSHVLWSRMMPNOFL KVDVSRISERGGVSSPEPPOLLKKTHMTLRIVANCALG NHRVAHALCSHVDEPGLLSVAYARAMMNNFYVFMT EETKSITLFPDENKKHGLPGGLSISSLRFRMGF]]	,					YGHAILLRHSYSGMYLCCLSTSRSSTDKLAFD
VGDDLLVSVSSERYLHLSYONGSLHVDAL QOTLWSVAPISSGRAGQVILGGDVLRLH GHMDECLTVPSGEHGEEQRRTVHYEGGAV VHARSLWRLETLRVAWSGSHRWGOPFALS HVTTGKYLSLMEDKNILLMDKEKADVKSTA FFFRSSKELDVGVRGEVDGMGTSEH YGGD VCVIGHVDTGLWLTYGSVDVSKSVMGSIGR KAMHHEGHMDDGISLSRSQHEESRTARVINGS TVPLFNRFIRGLDALSKKAASTVDLPIESVSL SLQDLIGYFHPPDEHLEHEDKONRLRALKNR QNLFGEGMINLVLECIDRLHVSSAHFAD VAGREAGESWKSILNSLYELLAALIRGNRKN CAOFSGSLDWLISHLERLEASSGILEVLHCVL VESPEALNIREGHKSIISLDICHGRNHKVLD VLCSLCVCHGYAVKSNGHLICDNLLPGRDLL LQTRLVNHVSSMRNIPLGVSEGSAQYKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYS YGGGEEWGGNOVGDDLFSYGFGGLHLWSG CLARTVSSPO;HLLRTDDVISCCLDLSAPSIS RINGOPVQGMEPSFNDLFFPVSFGGLHLWSG CLARTVSSPO;HLLRTDDVISCCLDLSAPSIS RINGOPVQGMEPSFNDLFFPVSFAGIKV RFLLGGRHGEFKFLPPFOVAPCVEAVLPKEKL KVEHSREYKQERTYTRDLLGFTVSLTOCH KVEHSREYKQERTYTRDLLGFTVSLTOCH KVEHSREYKQERTYTRDLLGFTVSLTOCH RELGWGYGFVGDDNCRGPCLVEFSKLPEQ ERNYNLQMSLETIKTLLALGGCHYGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLABENTALTLALGGCHYGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLABENHANWARABRGGWTY GIQQDVKNRNRPLVPYTFLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHARAPSCS LREAVRTLLGYGYNLEAPDQDHARAPSCS GTGEFFEIRAEKTYAVKAGRWYFEFETVTA GDMRYGWSSPGCQPQGLGSDERAFAFDGF KAQRWHCNSHYGRSWGAGDVYCGMVDM NEHTMAFTLNGEILLDDSGELAFKDFDVOG SKARKVLGSGENGRAFAFDGF KAQRWHCNSHYGRSWGAGDVYCGMVDM NEHTMAFTLNGEILLDDSGELAFKDFDVOG SINDIMFYRLSNPIECAEVFSKTVAGGLOAG LGGGYEPFAVNTNRDITMWLSKRLPGFLQA BRYDKGSKGRYCYSTRAHGHLVA DRVDKDKEATTFPNHKDYAGEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DRVDKDKEATTFPNHKDYAGEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DRVDKDKEATTFPNHKDYAGEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DRVDKDKEATTFPNHKDYAGEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DRVDKDKEATTFPNHKDYAGEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DRVDKDKEATTFPNHKDYAGEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DRVDKDKEATTFPNHKDYAGEFSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DRVDKDKEATTFPNHKDYAGEFSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DRVDKDKEATTFPNHKDYAGELFER RKSPCYYVCAGESMSFQCGRNNGLEIGE VVDAASGLLTFIANGKEISTYYOVFSTSLE BKSNCYYVCAGESMSFQCGRNNGLEIGE SKSSEVSTSNECOYOSPEPPIDLIKSKTIMPLIN EETKSITLFPDENKKHOLPEGLISLKTIM					Î		VGLQEDTTGEACWWT1HPASKORSEGEKVR
QQTLWSVAPISSGEAAQCYLIGDUNLLII GHMDECLTYPSGEHGEGRORTHYPGGAVS VHARSLWRLETLRVAWSGSHIRWGOPPELR HVTTGKYLSLMEDKYLLLMDKEKADVKSTA FTFRSSKEKLDVGVRKEVDGMGTSEIKYGDS VCYIGHVDTGLWLTYQSVDVKSVRMGSIGQ RAMHEGHMDGGSLSRSQHEESRTARVIRS TVFLFNRFIRGLDALSKKAKASTVDLPIESVS SLQDLIGYFHPPGEHLEHEDKORNRRALKNR QNLFQEEGMINLVLECIDRLHVYSSAAHFAD VAGREAGESWKSLINSLYBLAALIRGNRKN CAQFSGSLDWLISRLERLEASSGLEVLHCVL VESPFALNIIKEGHIKSIISLLDKHGRNHKVLDL JCSLCVCHGVAVRSNQHLICDNLLPGRDLL LQTRLVNHVSSMRPNIFLGVSEGSAQYKKW YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEGWGGNGVGDDLFSVGFDGLHLWSG CLARTYSSPNQHLLRTDDVISCCLDLSAPSISS RINGQPVQGMFENFINDGLFFPVVSFSAGIKV RFLLGGRHGEFKLPPGVAPCYEAVLPKGL KVEHSFEYKGERTYTRDLLGFTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENHELLWANN KIELGWYGFVRDDNKRQHPCLVFSKLPEQ ERNYNLQNSLETLKTLLALGCHVGISDEHAL DEVKKKMLPKNVQLTSGYKPAPDLSFRKL DEVKKMALPKNVQLTSGYKPAPDLSFRKL GDRAYNLGNSLETLKTLLALGCHVGISDEHAL DEVKKMALPKNVQLTSGYKPAPDLSFRKL GDRAYNLGNSLETLKTLLALGGHVGISDEHAL DEVKKMALPKNVQLTSGYKPAPDLSFRKL GDRAYNLGNGHARAREVCS GTGERFBIFRAEKTYAVKAGRWYFEFETVTA GDRAYGWSSPGCQPDQLLGSBERAFAFDGF KAQRWHGGNEHYGRSWQAGDVVGCMVDM NEHTMAFTLNGEILLDDGSGELAFRDFOG GFPVCSLGVAQVGRMPGKDVSTLKYFTIC GLQGGYEFFANTNRDITISMSLKRLPGPLOG GFPVCSLGVAQVGRMFGKDVSTLKYFTIC GLQGGYEFFANTNRDITISMSLKRLPGPLOG GFPVCSLGVAQVGRMFGKDVSTLKYFTIC GLQGGYEFFANTNRDITISMSLKRLPGPLOG GFPVCSLGVAQVGRMFGKDVSTLKYFTIC GLQGGYEFFANTNRDITISMSLKRLPGPLOG GFPVCSLGVAQVGRMFGKDVSTLKYFTIC GLQGGYEFFANTNRDITISMSLKRLPGPLOG GFPVCSLGVAQVGRMFGKDVSTLKYFTIC GLQGGYEFFANTNRDITISMSLKRLPGPLOG GFPVCSLGVAQVGRMFGKDVSTLKYFTIC GLQGGYEFFANTNRDITISMSLKKTLOGLLOG GFPVCSLGVAQVGRMFGKDVSTLKYFTIC GLQGGYEFFANTNRDITISMSLKSRLPGPLOG GFPVCSLGVAQVGRMFGKDVSTLKYFTIC GLQGGYEFFANTNRDTISMSLKSRLPGPLOG GFPVCSLGVAQVGRMFGKDVSTLKYFTIC GLQGGYEFFANTNRDTISMSLKSRLPGPLOG GFPVCSLGVAQVGRMFGKDVSTLKYFTIC GLQGGYEFFALTHAGGLGLGC VVDAASGLLTFANGKELSTYYQVEPSTLIPEN SICRSNCTYMCAGESMSPQGRNNNGLEIGC VVDAASGLLTFANGKELSTYYQVEPSTLIPEN GRANSTVDLETLGGELLKFHYTHLLYSAVCALG NHRVAHALGSHVDEPGLLKKTNVMPLIK EKTSILTFDENKKHGLPGGLGISTSLRFRMG GSFSVSISNEC(VGYSEPPPLOLLKKNTIONLTE	i		i				VGDDLILVSVSSERYLHLSYGNGSLHVDAAF
WHARSLWRLETLRVAWSGSHIRWGGPFRLR HVTTGKYLSLMEDKNLLIMDKEKADVKSTA FTRSSKEKLDVGVKEVDGMGTSEIKYGDS VCYQRIVDTIG WLTYQSVDVKSVMGSIQR KAIMHEGHMDDGISLSRSQHEESRTARVIRS TVELFFRERGLDALSKKAKASTVDLFESVSL SLQDLIGYTHPDEHLEHEDKONRLRALKIR QNLFQEEGMINLVLECIDRLHVYSSAAHFAD VAGREAGESWKSLINSLYLALIRGNRKM CAQFSGSLDWLISRLERLEASSGILEVLLGVL VESPFALNIKEGHKSIISLLDKHGRNHKVLD VLCSLCVCHGVAVRSNQHLICONLLFORDLL LQTRLVNHYSSMAPHILGVSGGSAQYKKWY YELMVDHTEPFVTAEATHLEVGASTGYSP YPGGGEWGGNGVGDDLFSYGFDGLHWSG CIARTVSSPNOHLLRTDDVISCLDLSAPSISF RINGQPVQGMEENFNIDGLFFPVVSFSAGIKV RFLLGGRHGEFKLPPPGVAPCYSEVLYFKKL KVEHSREYKQERTYTRDLLGPTVSLTOAAFT PIPVDTSQIVLPPHLERIEKLAENHELWWMN KEELGWYGYPYDDDNKRCHPCLVEFSKLPEQ ERMYNLQMSLETLKTLLALGCHYGISDEHAE DKVKKMKLPRNYQLTSGYKAPAMDLSFIKLT PSQEAMVDKLAENARINVWARDRIRQGWTY GIQQDVKNRRIPRLVPYTPLDDRTKKSNKDS LREAVRILLGYGYPLEAPDQDHARASCYC GTGERFIRFAEKTYAVKAGRWYFEFETVYA GIMPVGWRRIPRLVPYTPLDDRTKKSNKDS LREAVRILLGYGYPLEAPDQDHARASCYC GTGERFIRFRAEKTYAVKAGRWYFEFETVYA GMRVGWSRRSQCOPDQELGSDERAFAFDGF KAQRWHGNEHYGRSWQAGDVYGGMVDM NEHTMMFTLNGEILLDDSGSELAFKDPDVGD GFPVCSLGVAQVGRNAPICKADPSLQV PSNHEHEVTRIDGTIDSSPCLKVTQSSFGSQN SNTDIMFYRLSWFECAEVFSKTVAGGLPGAG LGGGYEFANTIRDITMUKSKLPPGLQV PSNHEHEVTRIDGTIDSSPCLKVTQSSFGSQN SNTDIMFYRLSWFECAEVFSKTVAGGLPGAG LGGGYEFANTIRRDITMUKSKLPPGLQV PSNHEHEVTRIDGTIDSSPCLKVTQSSFGSQN SNTDIMFYRLSWFECAEVFSKTVAGGLPGAG LGGGYEFANTIRRDITMUKSKLPGHLQV PSNHEHEVTRIDGTIDSSPCLKVTQSSFGSQN SNTDIMFYRLSWFECAEVFSKTVAGGLPGAG LGGGYEFANTIRRDITMUKSKLPGHLQV PSNHEHEVTRIDGTIDSSPCLKVTQSSFGSQN SNTDIMFYRLSWFECAEVFSKTVAGGLPGAG LGGGYEFANTIRRDITMUKSKLPGHLQV PSNHEHEVTRIDGTIDSSPCLKVTQSSFGSQN SNTDIMFYRLSWFECAEVFSKTVAGGLPGAG LGGYEFANTIRRDITMUKSKLPGHLQV PSNHEHEVTRIDGTIDSSPCLKVTQSSFGSQN SNTDIMFYRLSWFECAEVFSKTVAGGGPGAG LGGGYEFANDLEDVTRIPGTHAVAGRESTYTQVPETSTLIPF AVFAQATSPHVYGPLEGRIKNOMLEGC VVDAASGLITFIANGELSTTVYQVETSTLIPF AVFAQATSPHVGPLLKHCLYMFLERKRMGP SSPSYSISMEQUALVQCLDLYALSKTTOMLTE BETKSTILFPDENKKHOLPGIGLTSTLRRRMGE SSPSYSISMEQUALVQCLDLYALSKTTOMLTE BETKSTILFPDENKKHOLPGIGLTSTLRRRMGE							QQTLWSVAPISSGSEAAQGYLIGGDVLRLLH
HYTTGKYLSLMEDKNILLMDKEKADVKSTA FIFRSSKEKLDGVRKEVDGMGTSEIKYGDS VCYIQHVDTGLWLTYGSVDVKSVRMGSIQR KAIMHEGHMDDGISLSRSQHEESTARVIRS TVELFREREIDALSKKAKASTVDLPIESVSL SLQDLIGYFHPPDEHLEHEDKONRIRATARVIRS TVELFREREIDALSKKAKASTVDLPIESVSL SLQDLIGYFHPPDEHLEHEDKONRIRATARVIRS QNLFQEEGMINULVELCIBLKHYSSAAHFAD VAGREAGESWKSILNSLYELLAALIRGRIKN CAQFOSGLDWLISLERLEASSGILEVLHCVL VESPEALNIIKEGHIKSIISLLDKHGRNIKVLD VLCSLCVCHGVAVRSNQHLICONLLFGRDLL LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEWGGNGVODDLFSGGSAQYKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEWGGNGVODDLFSPOLGHLWSG CLARTVSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVQGMFENFNIDGLFFPVVSFRAGIKV RFLLGGRHGEFKFLPPPQYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGPTVSLTOAAFT PIPVDTSQIVLPPHLERIREKLAENHFELWVMN KEILGWYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETHATLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFRIK,PEQ ERNYNLQMSLETHATLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFRIK,PEQ ERNYNLQMSLETHATLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFRIK,PEQ ERNYNLQMSLETHATLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFRIK,PEQ GROUPKNRRNPRLYVTTLDDRTKKSNKDS LREAWTILLGYGYNLEAPDQDHAARAEVCS GTGEFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRGCOPDQELGSDERAFAFDGF KAQRWHQONEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSLAFKDPDVGD GFIPVCSLGVAQVGGRMFGKDVSTLKYTTIC GLQCYEFFANTINDITINVLSKLPQDLQV PSNIEHHEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMFECABVESTLVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKRATKPETNNIKDYAQEKPSRIKQ RFLLRRTKPDYSTSISSALLEDULADDRDDY DFLMQTTSTYYSVSIFPGGEPANVWCGMTS DFHQVDTGFDLDRYRTVTVTLOEGEGKYHE DFHQVDTGFDLDRYRTVTVTLOEGEGKYHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGELSTYYQVERSTKLPP AVFAQATSPNWYPGELGRIKNYMFLSAGLFKS EHKNPYQCPPRLHVQFLSHVLWSRMPQFL KVDVSRISERQGVLVQCLDPCJPMSLHIPEEN RSVDLELTEQEELLKHYHTILLYSAVCALG NIRWAHALCSHVDEPQLLYSARNMFEYNPMT BETKSTILFPDENKKHOLPGIGLSTSLRRRMQF SSPSYSISNECYGYSPEPPLULRSKTIOMLTE BETKSTILFPDENKKHOLPGIGLTSTSLRRRMGP			ł				GHMDECLTVPSGEHGEEORRTVHYEGGAVS
FTFRSSKEKLDVGVKEVDOMGTSEKYGDS VCYIQHVDTGLWLTYQSVDXSVRMGSIQR KAIMHEGHMDDGISLSRSQHEESRTARVIRS TVFLPNFRIRGLDALSKKAKASTVDLPESVSL SLQDLIGYTHPPDEHLEHEDKOPKIRALKNR QNLFQEEGMINLVLECIDRLHVYSSAAHFAD VAGREAGESWSKINNSYLELALIRGNRKAN CAOFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHIKSISLLDKHGRNHKVLD VLCSLCVCHGVAVRSNQHLICONLLPGRDLL LQTRLVNHVSSMAPHIRLGVSEGSAQVKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEWGGNGVGDDLFSYGFDGLHLWSG CLARTVSSPNQHLLRTDDVISCCLDLSAFSISF RINGQPVQGMEENPNIDGLFFPVVSFAGIKV RFLLGGRHGEFKFLPPFQYAPCYEAVLPKEKL KVEHSREYXGRETYTRDLDGTFPVVSFAGIKV RFLLGGRHGEFKFLPPFQYAPCYEAVLPKEKL KVEHSREYXGRETYTRDLDGTCHVSLTQAAFT PIPVDTSQIVLPPHLERREKLAENHFELWVMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENASHNVARAFCS GTGERFRIFRAEKTVAVKAGRWYFEFETVTA GOMDVKNRRIPRLVPYTFLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARACVCS GTGERFRIFRAEKTVAVKAGRWYFEFETVTA GOMRVGWSRGCQPDQELGSDERAFADGF KAQRWHGGNEHYGRSWQAGRWYFEFETVTA GOMRVGWSRGCQPDQELGSDERAFADGF KAQRWHGGNEHYGRSWQAGRWYFEFETVTA GOMRVGWSRGCQPDQELGSDERAFADGF KAQRWHGHFYGRSWQAGRWYFEFETVTA GOMRVGWSRGCQPDQELGSDERAFADGF KAQRWHGHFYGRSWQAGRWYFEFETVTA GOMRVGWSRGCQPDQELGSDERAFADGF KAQRWHGHFYGRSWQAGRWYFEFETVTA GOMRVGWSRGCQPDVGCGMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFFPVCSLGVAQVGRMYFGKDVSTLKYFTIC GLQEGYEFFAWNTNDITMWLSKRLPQFLQV PSNIEHIEVTRDGTIDSSPCLKVTQKSFGSQN SNITDIMFYRLSMPECAEVSKTVAGGLPGAG LFGFKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKAATKFEFNNHKDYAQGLFGKWHTS DFHQVTOTGFDLDRVTNTVTHLGDEGKGVHE SIKRSNCYMVCAGESMSPQQGRNNGLGICC VVDAASGLLTFLINGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRINVMFLSAGLFKS EHKNPVQCPPRLHVQFLSHVLWSRMPOQFL KVDVSRISERQGVLVQCLDPLCFMSLHIFEEN RSVDLELTERGELLKHYHTLRLYSAVCALG NIRVAHALCSHLFPDNIKHGLYSTHMFEN RSVDLELTERGELLKHYHTLRLYSAVCALG NIRVAHALCSHLFFPNIKHGLIGSTSLRRMQFI RVDVSRISERQGVLVQCLDPLJCSTSLRRRMQFI RSVDLELTEGDELLKHYHTLRLYSAVCALG NIRVAHALCSHLFPDNIKHGLIGSTSLRRRMQF SSPSFYSISNECYQYSPEPULDLISKTUMLE EKTSSITTFDDEKKHGLPGIGLSTSLRRRMQF SSPSFYSISNECYQYSPEPULDLISKTUMLE	1	1					VHARSLWRLETLRVAWSGSHIRWGQPFRLR
VCYIQHVDTGLWLTYQSVDVKSVKMGSIQR KAMHHEGHNDDGISLSRSQHEESRTARVIRS TVFLFNFRGLDALSKKAKASTVDLPIESVSL SLQDLIQYTHPDEHLEHEDKQNRLAKINR QNLFQEEGMINLVLECIDRLHVYSSAAHFAD VAGREAGESWKSILNSLYELLAALIRGNRKN CAQFSGSLDWLSILERLEAGIBEVLHCVL, VESPRALNIKEGHIKSIISLLDKHGRNHKVLD VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL LQTRLVNHVSSMRPNIFLGVSGSAQYKKWY YELMVDHTEFFVTAATHILRVGWASTEGYSP YPGGGEWGGNGVDDLFSYGFDGLHLWSG CIARTVSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVQGMEENNFIDGLFSVGFDGLHLWSG CIARTVSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVQGMEENNFIDGLFSVGFDGLHLWSG CIARTVSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVQGMEENNFIDGLFSVGFSAGIKV RFLLGGRHGEFKFLPPPQYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGPTVSLTOAAFT PIPVDTSQIVLPPHLERIREKLAENHFELWYMN KELGWQYGPVRDDNKRQHPCLVEFSKLPQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGTYY GIQQDVKNRRNPRLVPYTHDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGTWFFEFTUTA GDMRVGWSRGCQPDQELGSDERAFAAFDGF KAQRWHOGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDPDVGD GFFVCSLGVAQVGRMFGKDVSTLKYTTIC GLQCGYEFFANTINDITINDITINVLSKRLPOPLQV PSNIEHIEVTRIDGTIDSSPCLKVTQSSFGQN SNITDIMFYRLSMPECAEVESKTYAGGLBGAG LFGPKDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKAATKPETNINKDYAQEKPSRIKQ RFLLRRTKPDYSTSISSALTEDVLADDRDDY DFLMQTTSTYYSVRIFFGQFPANVWGMTS DFHQVDTGTYSTSISSALTEDVLADDRDDY DFLMGTSTYYSVRIFFGGGFANWWGMTS DFHQVDTGTPLDRYRTTVTLDGEGKVHE SIKRSNCYMVCAGESMSPQGGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEFSTKLPF AVFAQATSPNNYQFELGRLKYNMFLSAGLFKS EHKNPYQCPPRLHVQFLSHVLWSRMPQGL KVDVSRISERQGVLVQCLDPLOFMSLHIFEEN RSVDLELTEQBELLKHYHTILLYSAVCALG NIRVAHALCSHVDEFQLLYAIENKYMPGLLR AGYYDLLDIHLSSYTAARMMNEYTYPMT EETKSTILFPDENKKHOLPGIGLSTSLRRRMQF SSPSYSISNECQYSPEPPLULRSKTIOMLE		1	i				HVTTGKYLSLMEDKNLLLMDKEKADVKSTA
KAIMHEGHMDDGISLSRSQHEESRTARVIRS TVFLFNRFIRGLDALSKAASTVDLPISVSL SLQDLIGYFHPPDEHLEHEDKQNRLRALKNR QNLFQEEGMINLVLECIDRLHVYSSAAHFAD VAGREAGESWKSILNSLYSLLAALIRGNRKN CAQFSGSLDWLISRLERLEASSGICHVLHCVL VESPEALNIRKEGHIKSISLLDKHGRNHKVLD VLCSLCVCHGVAVRSNOHLICDNLLPGRDLL LQTRLVNHVSSMRPNIFLGVSEGSAGVKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEEWGGNGVGDDLFSYGFOGHLH WSG CIARTVSSPROHLLRTDDVISCCLDLSAPSISF RINGOPVQGMFESHNIDGLFFPVVSFSAGIKV RFLLGGRHGEFKFLPPPGYAFCYALPKEKL KVEHSREYKQERTYTRDLLGPTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENHELWVMN KIELGWQYGPVEDDNKRGHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKAPPMDLSIKLT PSQEAMVDKLAENAHNWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERRIFRAEKTYAVKAGRWYFEFETVTA GDDMKVGWSRPGCPDQELGSDERAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTIMMFTLNGEHLLDGSGSLARAFDGF KAQRWHQGNEHYGRSWAGDVVGCWDDM NEHTIMMFTLNGEHLDDSGSSLARAFDGF GFFPVCSLGWAQVGRMNFGKDVSTLKYTFIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHEVTRUGDTIDSSTCLKVTGKSFGSON SNTDMFYKLSMPECAEVTSKTVAGGLPGAG LFGFRNDLEDYDADSDFEVLMKTAGHLVP DRVDKDKEATKPFFNNHENDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTTYYYSVHFPGGPANVWGWTTS DFHQYDTGFDLDRVRTVYTVLGDEKGKVHE SIKRSNCYMVCAGESMSFGQGRNNNGLEIGC VVDAASGLLTFANGKLETSHVAYGQFRSTLLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSRLLFPHAVFAGLLFR RSVDLLELTEQEELLKFHYHTLRLYSAVCALG NIRKYABALGSTUPPEDNKKHGLPGILSTSLRFMQF RSVDLLELTEQEELLKFHYHTLRLYSAVCALG NIRKYABALGSHUPGLIRKSKTIQMLIPE RSVDLLELTEQEELLKFHYHTLRLYSAVCALG NIRKYABALGSHUPGLIRKSKTIQMLFB							FTFRSSKEKLDVGVRKEVDGMGTSEIKYGDS
TVFLFNRFIRGLDALSKKAKASTVDLPIESVSL SLODLIGYPHPPDEHL EINDONDRIALKNR ONLFQEEGMINLVLECIDRLHVYSSAAHFAD VAGREAGESWKSILNSLYELLAALIRGNRKN CAQFSGLDVUISILERLSSGILEVLHCVL VESPEALNIIKEGHIKSIISLLDKHGRNIKKVLD VLCSLCVCHGVAVRSNOHLICDNLLPGRDLL LQTRLVNHVSSMRNHICJOSEGSAQYKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSP PYPGGEEWGGNGVGDDLSAPSISF RINGQPVQGMFENFINDGLFFVVSFSAGIKV RFLLGGRHGEFKFLPPPGYAPCVEAVLPKEKL KVEHISREYKQERTYTRDLLGPTVSFSAGIKV KVEHISREYKQERTYTRDLLGPTVSTSAGIKV RFLLGGRHGEFKFLPPPGYAPCVEAVLPKEKL KVEHISREYKQERTYTRDLLGPTVSTSAGIKV NELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNWARDRIRQGWTY GIQQDVKINRRNPRLYPYTDDATKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERREFRAEKTYAVKAGRWYFEFTVTA GDMRVGWSRPGCQPDQELGSDERAFAPDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM MEHTMMFTLNGEILLDDGSGLAFKDFDVGD GFBVCSLGWAQVGRMNGKDDSTLKYTTIC GLQEGYEPFAVNTNRDITMWLSKRIPQFLQV PSNHEHIEVTRIDGTIDSSELAFKDFDVGD GFBVCSLGWAQVGRMNGKDDSTLKYTTIC GLQEGYEPFAVNTNRDITMWLSKRIPQFLQV PSNHEHIEVTRIDGTIDSSELAVTKOFLOGAG LFGPKNDLEDYDABSDFEVLMKTAGHLVP DRVDKDKEATKYPFNNHENDYAQEKPSRLKQ RFILARTKPDYSTSHSARLTEDVLADRODY DFLMGTSTTYYSYRIFPGQEPANVWVGWITTS DFHQYDTGFDLDRVRTVTVTLODEKGKVHE SIKRSNCYMVCAGESMSPGQCRNNNGLEIGC VVDAASGLLTFANGKELSTLAPDRODY DFLKNTTYYSYRIFPGGPEANVWVGWITTS DFHQYDTGFDLDRVRTVTVTLODEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFANGKELSTLAPHDFNSLERFMOFT RSVDLLELTEQEELLKFHYHTLRLYSAVCALG NIRKPAHALGSHUKHGLPGIGLSTSLRPMOFT RSVSDLLELTEQEELLKFHYHTLRLYSAVCALG NIRKPAHALGSHUKHGLPGIGLSTSLRPMOFT RSVSPLELITEQEELLKFHYHTLRLYSAVCALG NIRKPAHALGSHUKHGLPGIGLSTSLRPMOFT RSVSPLELITEQEELLKFHYHTLRLYSAVCALG NIRKPAHALGSHUKHGLPGIGLSTSLRPMOFT RSSPSFYSISNECYQVSSPFFIDLIKSKTIOMLTE		- 1		j	}		VCYIQHVDTGLWLTYQSVDVKSVRMGSIQR
SLQDLIGYFHPPDEHLEHEDKQNRLRALKNR QNLFQEEGMINLVLECULHYVSSAAHFAD VAGREAGESWKSILNSLYELLAALIRGNRKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIKEGHIKSISILDKHGRNIKVLD VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL LQTRLVANHVSSMRPHIFLOXEGSAQVKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEEWGGNGVGDDLFSYGFOIGHLWSG CIARTVSSPROHLLITDDVISCCLDLSAPSISE RINGQPVQGMFEYNIDGLFFPVVSFSAGIKV RFILLGGRHGEFKPLPPDYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGFTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENHELWYMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERMYNLQMSLETLKTLLALGCHYGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLYPYTTDDDRTKKSNKDS LREAVRTILLGYGNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDDMEYGWSPGCQPDQELGSDERAFAFDGF KAQRWHQONEHYGRSWQAGDVVGCMYDM NEHTMMFTLNGEHLLDSGSELAFKDFDVGD GFFPVCSLGWAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDMFYRLSMPCAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPFFNNHAYGHENJAGHLVP DRVDKDKEATKPFFNNHAYGERPSRIKQ RFLLARTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFFGQEPANWVGWTTS DFHQYDTGFDLDRVRTVYTLGDEKGKVHE SIKRSNCYMVCAGESMSFGQGRINNDLEIGC VVDAASGLLTFANGKLESTYYQVEPSTKLPP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQLISHVLWSRMPOGLR RSVDILELTEQEELLKFHYHTLRLYSAVCALG NIRKVAHALGSTUPPENDLKSTTIQMLTE		ļ	i	İ	1		KAIMHHEGHMDDGISLSRSQHEESRTARVIRS
QNLFQEEGMINI.VLECIDR.HYYSSAAHFAD VAGREAGESWKSII.NSLYELIAALIRGNRKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALMIIKEGHIKSIISLIDKIGRNIKVLD VLCSLCVCHGVAVRSNQHLICDNLIPGRDLL LQTRLVNIHYSSMRPNIFLGVSEGSAQYKKWY YELMYDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEEWGGNGVGDDLFSYGFDGLHLWSG CLARTYSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVQGMFENFINDGLFFPVVSFSAGIKV RFLLGGRHGEFKFLPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGFTPVSFSAGIKV RFLLGGRHGEFKFLPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGFTVSLTQAAFT PIPVDTSQIVLPPHLERIEKLAENHELWYMN KIELGWQYGPVRDDNKRQHFCLVEFSKLPEQ ERNYNLOMSLETLKTLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMYDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTDDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYEFETTYTA GDMRVGWSRRGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGGLAFAFDGF KAQRWHQGNEHYGRSWQAGGLAFAFDGF KAQRWHQGNEHYGRSWQAGGLAFAFDGF GAQRWHGWSRGCQPDGLGSDERAFAFDGF GAQRWHGWSRGCQPDGLGSDERAFAFDGF GAQRWHGWSRGCQPDGLGSDERAFAFDGF KAQRWHGNEHYGRSWQAGGLAFAFDGF GARWHTTLNGEILDSSPCLKVTOKSFGSON SNTDIMFYLLSRPECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVY PRAVBELTYTRYSTISSSCLKVTOKSFGSON SNTDIMFYRLSRPECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVY PRAVBAGATTYFTSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVEPGEAPNWWQWITS DFHQYDTGFDLDRVRIVTVTLGDEKGKVHE SKRSNCYYWCAGESMSPQGGRNNIGLEIGC VVDAASGLITFHANGELSTTYYQVEPSTKLPP AVFAQATSPNYFQFELGRIKNWMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVWRSMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVHAHACLSHVUEPGLLYAENKYMPGLIR AGYYDLLIDHLSSYATARLMMNNEYVPMT EETKSTILFPDENKKHGLPGIGLSSTLRPRMQF	ľ	}					TVFLFNRFIRGLDALSKKAKASTVDLPIESVSL
VAGREAGESWKSILNSLYELLALIRGNRKN CAQFSGSLDWLISRLERLEASSGILEVLHCVL VESPEALNIIKEGHIKSIISLLDKHGRNIKVLD VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL LQTRLVNIVSSMRPNIFLGVSGSAGVKKWY YELMVDHTEFFVTAEATHLRVGWASTEGYSP YPGGGEWGGNGVGDDLFSVGFDGLHLWSG GLARTVSSRNQHLLRTDVISCCLDLSAPSISF RINGQPVQGMFENFNIDGLFFVVSFSAGIKV RFLLGGRIGEFKFLPPPGVAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGFTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENHELWVMN KIELGWQGPVRDDNKGPHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCOPDQELGSDERAFAFDGF KAQRWYGNSPGCOPDQELGSDERAFAFDGF KAQRWYGNSPGCOPDQELGSDERAFAFDGF KAQRWYHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILDDSGSELAFKDFDVGD GFFPVCSLGVAQVGRMFGKDVSTLKYTTIC GLQGGYGFPAVNTNADITMWLSKRLPOFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFOFKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATTFEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIPPGGEPANVWGWITS DFHQYDTGFDLDRVRITVTVLIGBEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNYFQFELGRIKNYMLSAGLFKS EHKNPVPQCPPLHAVDFLSHVLWSRMPNQFI KVDVSRISERQWLVQCLDPLOFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRYSAVCALG NHRVAHALCSHVUEPGQLIYAERKYMPGLIR AGYYDLLDDHLSSYATARLMMNNEYIVPMT EETKSITLFPDEKKHGLPGIGLSSTLRPRMQF		ł	- 1	I			SLQULIGYFHPPDEHLEHEDKQNRLRALKNR
CAQFSGSLDWLISELERLEASSGILEVLHCVL VESPEALMIKEGHIKEJCHNEKVLD VLCSLCVCHGVAVRSNQHLICDNLLPGRDLL LQTRLVNIHVSSMRPNIFLGVSEGSAQYKKWY YELMYDHTEPFVTARATHLRVGWASTEGYSP YPGGGEEWGGNGVGDDLFSVGFDGLHLWSG CLARTVSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVQGMFENFINDGLFFPVVSFSAGIKV RFLLGGRIGEFKFLPPPGVAPCVEAVLPKEKL KVEHSREYKQERLYTDLLGPTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENIELLWVMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEG ERNYNLQMSLETLKTLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVARARRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGEFREFREAKETYA VKAGRWYFEFFTVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDWGCMVDM MEHTMMFTLNGEILIDDSGSELAFVDFVGD GFPVCSLGVAQVGRMNFGKDVSTLKYPTIC GLQEGYEPFAVNTNRDITMVLSRLPQFILVQV PSNHEHIEVTRDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDEVLMKTAHGHLVP DRVDKDKEATKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVEIPFGGEPANVWGWTTS DFFIQYDTGFDLDRVRTVTTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNGLEIGC VVDAASGLLTFIANGKELSTTYYQVEPESTKLPP AVFAQATSPNYFQFELGRIKNVMPLSAGLFKS EHKNPYPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLOFRISLHPEEN RSVDILELTEQEELLKFHYHTLRYSAVCALG NHRVAHALCSHVDEPQLLYAERKYMPGLR AGYYDLLDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTLRPRNQF		1					VAGREA GESTIVET NOT WELL AND TO THE
VESPEALNIKEGHIKSIISLIDKHGRNHKVLD VLCSLCVCHGVAVRSNQHLCDNLLPGDLL LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY YELMVDHTEFFVTAEATHLRVGWASTEGYSP YPGGGEWGGNOGDDLFSYGFDGJEHLWSG CLARTVSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVQGMFENFINDGLFFVVSFSAGIKV RRLLGGRHGFKFLPPGVAPCYFAVLPKEKL KVEHSREYKQERTYTRDLLGPTVSLTQAAFT PIPVDTSQIVLPPHLERIFEKLAENHELWVMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DEVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDIRQGWTY GIQQDVKNRRNRLVPYTELDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQONEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDOF GFIPVCSLGVAQVGRMFNGKDVSTLKYFTIC GLGEGYEFFAVNTNRDITMVLSKRLPQFLQV PSNIEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKVAGGLPGAG LFGFKNDLEDVDADSDFEVLMKTAHGHLVP DRVDKDKEATKPPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIPFQGPANVWVGWITS DFHQYDTGFDLDRVRTYTVTLGBEKGKVHE SIKRSNCYMCCAGESMSPCQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLEP AVFAQATSPNVFQFELGRIKNYMFLSAGLFKS EHKNPVPCQPPRLHVFLSHVLWSSWPNOGL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEGEELLKFHYTHTLLYSAVCALG MHRVAHALCSHVDEPQLLVAIENKYMPGLER AGYYDLLIDHILSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLISTSLRPRMQF SSPSYVSISNECYQVSPEFFLDLKSKTIQMLTE			1	ĺ	1		CACEGGI DWI ISBI EDI EASSOII DIE HOUR
VI.CSI.CVCHGVAVRSNGHLICDNILPGRDLL LOTRLVNHVSSMRPNIFLGVSEGSAQYKKWY YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGEEWGGNGVGDDLFSYGFDGLHLWSG CLARTVSSPNOHLLRTDDVISCCLDLSAPSISF RINGQPVQGMFENFNIDGLFFPVYSFSAGIKV RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL KVEHSREYXGGETYTRDLLGFTVSLTQAAFT PIPVDTSQIVLPPHLERIEKLAENHELWVMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETIKLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLYPTPLDDRTKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSLAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKAKATKPEFFNNHKDYAGEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTTYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGOGRNNINGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQELGRIKNVMFLSAGLFKS EHKNPVPQCPPRLHRVFLSHVUSSHNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQELGRIKNVMFLSAGLFKS EHKNPVPQCPPRLHVOFLSHVLWSSRMPNOFL KVDVSRISERGGWLVQCLDPLOFMSLHIPEEN RSVDILELTEQEELLKFHYHTIRLTYSAVCALG MHRVAHALCSHVDEPQLLVAIENKYMPGLLR AGYYDLLIDHILSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHOLDGIGLSTSLRPRMQF			l				VESDEVI MILLEGRINGRICI I DATICUMITATA D
LQTRLVNHVSSMRPNIFLGVSEGSAQYKKWY YELMVDHTEPFVAEATHLRVGWASTEGYSP YPGGGEWGGNGVGDDLFSYGFDGLHLWSG CLARTVSSPNOHLLRTDDVISCLDLSAPSISF RNGGPVQGMFSYFNIDGLFFPVSSCAGIKV RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLDFTVSLTQAAFT PIPVDTSQIVLPPH_ERIREKL_AENIEL_WVMN KIELGWQYGPVRDDNKRQHPCL_VEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYGTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLYPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGEFRIFFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHGONEHYGRSWQAGDVVGCMYDM NEHTMMFTLNGEILLDDSGSLAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMVLSKRLPQFLQV PSNIEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDDMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHL.VP DRVDKDKEATKFEPNYSTRABLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVVVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMCCAGESMSFQGRNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVPGFELGRIKNVMFLSAGLFKS EKNPVPCQCPPRLIKVFINIGLISGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVPGFELGRIKNVMFLSAGLFKS EKNPVPCQCPPSLHVFLSRWPNOGL. KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEGEELLKFHYHTILLYSAVCALG MHRVAHALCSHVDEPQLL YAIENKYMPGLER AGYYDLLIDHILSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF			į				A CAL CACHERA A DESTRUCTION TO THE ACTION OF
YELMVDHTEPFVTAEATHLRVGWASTEGYSP YPGGGEEWGGNDCPSYGFDGLHLWSG CIARTVSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVQGMFENFNIDGLFFPVVSFSAGIKV RFLLGGRHGFFKFLPPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGPTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENHELLWVMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVSTILGYGYNLEAPDQDHAARAEVCS GTGERRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAPTDGF KAQRWHQGNEHYGRSWQAGDDVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQGGYEPFAVNTNRDITMWLSKRLPGFLQV PSNIEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SINTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGFKNDLEDYDADSDFEVLMKTAHGFILVP DRVDKDKEATKPEPNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTTYYSVRIFPGQEPANVWVGWTTS DFHQVDTGFDLDRVTNTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQCFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLVOFLSHLWSRMPNOFL KVDVSRISERQGWLVQCLDPLOFMSLHIPEEN RSVDILELTEQEELKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENFYMPGLLR AGYYDLLIDHLSSVATARLMMNNEYIVPMT EETKSTILFPDENKKHQLIJKAIENFRMQF	1	Í	- 1	1	- 1		LOTRI VNHVSSMR PNIFI GVSEGS A OVENUA
YPGGGEEWGGNGVGDDLFSYGFDGLHLWSG CIARTVSSPNQHLLRTDDVISCCLDLSAPSISF RINGQPVQGMFENFNIDGLFFPVVSFSAGIKV RFLLGGRHGFKFLPPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGPTVSLTQAAFT PIPVDTSGIVLPPHLERIREKLAENHELWVMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLAGCHYGISDEHAE DKVKKMLPKNYQLTSGYKPAPMDLSFRLY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGGYNLEAPPQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEVGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSLAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNIEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNIDIMFYRLSMPECAEVTSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPFNNHKDYAGERPSRLKQ RFLLRTKFDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGGEPANVWGWTTS DFHQVDTGFDLDKTVTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFOFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNOFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDIL ELTEGEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYALENKYMPGLLR AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSTILFPDENKKHQLPGIGLSTSLRPRMQF							YELMVDHTEPFVTAFATHI RVGWASTEGVER
CIARTYSSPNQHLISTDOVISCCLDLSAPSISF RINGQPVQGMFENFNIDGLFFPVVSFSAGIKV RFLLGGRHIGEFKFLPPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGPTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENIHEL WVMN KEELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHYGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNIVWARDRIRQGWTY GIQQDVKNRRNPRLYPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRYGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEHLDDSSGELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMVLSKRLPQFLQV PSNHEHEVTRIDGTIDSSPCLKVTQKSFGSON SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFINNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTTSTYYYSVRIFPGGEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPQGRNNNGLEIGC VVDAASGLITFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFGELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNOFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAERKYMPGLR AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGISTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	1	1		1	. [YPGGGEEWGGNGVGDDI FSVGFDGI HI WSG
RINGQPVQGMFENFINDGLFFPVVSFSAGIKV RFLLGGRHGEFKPLPPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGPTYSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENIHELWVMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKITLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHINVWARDRIRQGWTY GIQQDVKNRNPRLYPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTINRDITMWLSKILPGFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDVDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKBSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHFPEEN RSVDILELTEQEELLKFHYHTRLLYSAVCALQ NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTB	ĺ	ĺ	- 1	ľ	- 1		CIARTVSSPNOHLLRTDDVISCOLDI SAPSISE
RFLLGRHGEFKFLPPPGYAPCYEAVLPKEKL KVEHSREYKQERTYTRDLLGPTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENHELWVMN KEELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLYPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTIDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNIKDYAQEKPSRLKQ RFLLRTKRPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGGEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPQQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQEELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLCPRMSLHIPEEN RSVDILELTEGELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNEKKHGLPGIGISTSLRPRMQF		1	Í	- }			RINGOPVOGMFENFNIDGLFFPVVSFSAGIKV
KVEHSREYKOERITYTRDLLGFTVSLTQAAFT PIPVDTSQIVLPPHLERIREKLAENIHELWVMN KIELGWQGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMLPKNYQLTSGYKPAPMDLSFIKLT PSOEAMVDLAENAHNVWARDRIRQGWTY GIQQDVKNRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILDDSGSELAFKDPDVGD GFFPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGGEPANVWGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNIGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFOFELGRIKNVMPLSAGLFKS EIKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLOFMSLHIPEEN RSVDLELTEQEELLKFHYHTIRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLIR AGYYDLLIDHILSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMOF				1	i		RFLLGGRHGEFKFLPPPGYAPCYEAVLPKEKI
PIPVDTSQIVLPPHLERIREKLAENIHELWYMN KIELGWQYGPVRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFFPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVTSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFILRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDHILSSYATARIMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSIVSISNECYQYSPEPIDILKSKTIOMLTE		ſ	ľ	1	.		KVEHSREYKQERTYTRDLLGPTVSLTOAAFT
KIELGWQYGPYRDDNKRQHPCLVEFSKLPEQ ERNYNLQMSLETLKTLLALGCHVGISDEHAE DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTINRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEIFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDHILSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSIVSISNECVQYSPEPPLDILKSKTIQMLTE	1	1	- 1	1	j		PIPVDTSQIVLPPHLERIREKLAENIHELWVMN
DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMAFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEFFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTINDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQVDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMFLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSSISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE		İ		İ			KIELGWQYGPVRDDNKRQHPCLVEFSKLPEO
PSQEAMVDKLAENAHNVWARDRIRQGWTY GIQQDVKNRRNPRLVPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYPTIC GLQEGYEFFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSSTLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE		1	- {				ERNYNLQMSLETLKTLLALGCHVGISDEHAE
GİQQDVKNRRNPRL VPYTPLDDRTKKSNKDS LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGEFRFIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFFLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	1	1	- 1	1	1)	DKVKKMKLPKNYQLTSGYKPAPMDLSFIKLT
LREAVRTLLGYGYNLEAPDQDHAARAEVCS GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPETNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYALENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	-			1			PSQEAMVDKLAENAHNVWARDRIRQGWTY
GTGERFRIFRAEKTYAVKAGRWYFEFETVTA GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSTTLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE				1			GIQUDVKNRRNPRLVPYTPLDDRTKKSNKDS
GDMRVGWSRPGCQPDQELGSDERAFAFDGF KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYPTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	1		ł	!		i	GTGERER A RETTY ALIV A CRUIT TO THE
KAQRWHQGNEHYGRSWQAGDVVGCMVDM NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYPTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNITDIMFYRLSMPIECAEVTSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE		1	- 1	i	1		GDMPVGWSPBGCOPDOEL CODED A FARDOR
NEHTMMFTLNGEILLDDSGSELAFKDFDVGD GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE						I	K V D M TO COLD COLD COLD COLD COLD COLD COLD COL
GFIPVCSLGVAQVGRMNFGKDVSTLKYFTIC GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHILVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	- 1	}	1	- 1	1	İ	NEHTMMETI NGEILL DOGGEL AEKDONGO
GLQEGYEPFAVNTNRDITMWLSKRLPQFLQV PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	j			1	İ		GFIPVCSLGVAOVGRMVFGKDVSTI KVETIC
PSNHEHIEVTRIDGTIDSSPCLKVTQKSFGSQN SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	1	1	- 1	İ	1		GLOEGYEPFAVNTNRDITMWI SKRI ROPE OV
SNTDIMFYRLSMPIECAEVFSKTVAGGLPGAG LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE			- 1	-	1		PSNHEHIEVTRIDGTIDSSPCI KVTOKSEGSOM
LFGPKNDLEDYDADSDFEVLMKTAHGHLVP DRVDKDKEATKPEFFNNHKDYAQEKPSRLKQ RFLLRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	. [j	.].].	. 1.	. 1	SNTDIMFYRLSMPIECAEVFSKTVAGGI PGAG
DRVDKDKEATKPEFNNHKDYAQEKPSRLKQ RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYSVRIFPGQEFANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	- 1	-	ĺ		1.	•	LFGPKNDLEDYDADSDFEVLMKTAHGHI VP
RFLLRRTKPDYSTSHSARLTEDVLADDRDDY DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMFLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE					1	ļ	DRVDKDKEATKPEFNNHKDYAOEKPSRI KO
DFLMQTSTYYYSVRIFPGQEPANVWVGWITS DFHQYDTGFDLDRVRTVTVLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE					1		RFLLRRTKPDYSTSHSARLTEDVLADDRDDY
DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE			1	1	ĺ	1	DFLMQTSTYYYSVRIFPGQEPANVWVGWITS
SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE		ĺ		1		1	DFHQYDTGFDLDRVRTVTVTLGDEKGKVHE
VVDAASGLLTFIANGKELSTYYQVEPSTKLFP AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE			Ì		i	ŀ	SIKRSNCYMVCAGESMSPGQGRNNNGLEIGC
AVFAQATSPNVFQFELGRIKNVMPLSAGLFKS EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	- 1	1				Ì	VVDAASGLLTFIANGKELSTYYOVEPSTKLFP
KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	1	}	- 1	1			AVFAQATSPNVFOFELGRIKNVMPLSAGLFKS
RSVDILELTEQEELLKFHYHTLRLYSAVCALG NHRVAHALCSHVDEPQLLYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE		1	Ì			[EHKNPVPQCPPRLHVQFLSHVLWSRMPNQFL
NHRVAHALCSHVDEPQILIYAIENKYMPGLLR AGYYDLLIDIHLSSYATARLMMNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	1		1		Į		KVDVSRISERQGWLVQCLDPLQFMSLHIPEEN
AGYYDLLIDIHLSSYATARLMMNNEYIVPMT EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE	J	- 1	1	- 1	j	1.	RSVDILELTEQEELLKFHYHTLRLYSAVCALG
EETKSITLFPDENKKHGLPGIGLSTSLRPRMQF SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE		-	.	1	[Į.	NHKVAHALCSHVDEPQLLYAIENKYMPGLLR
SSPSFVSISNECYQYSPEFPLDILKSKTIQMLTE			Ì			į.	AGYYDLLIDIHLSSYATARLMMNNEYIVPMT
SSYST VSISNECYQYSPEFPLDILKSKTIQMLTE AVKEGSI HARDPVGGTTFFI FVDI IKI EVTI I I			J	j	ı		BEREIT OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF THE PROPERTY OF T
——————————————————————————————————————	1			Ì	[ĺ	AVECE HAPPING CTTTT TO THE
[···· · · · · · · · · · · · · · · · ·							AVAEUSLIAKUPVUUTTEFLFVPLIKLFYTLLI

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D-Aspartic Acid. E-Glutamic Acid.
nuci-	peptide	i	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	}	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	Ì	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	i	į		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
	1		l	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						MGIFHNEDLKHILQLIEPSVFKEAATPEEESDT
		•				LEKELSVDDAKLQGAGEEEAKGGKRPKEGLL
						QMKLPEPVKLQMCLLLQYLCDCQVRHRIEAI
1		l . i				VAFSDDFVAKLQDNQRFRYNEVMQALNMSA
1						ALTARKTKEFRSPPQEQINMLLNFKDDKSECP
						CPEEIRDQLLDFHEDLMTHCGIELDEDGSLDG
						NSDLTIRGRLLSLVEKVTYLKKKQAEKPVES
						DSKKSSTLQQLISETMVRWAQESVIEDPELVR
				ĺ		AMFVLLHRQYDGIGGLVRALPKTYTINGVSV
						EDTINLLASLGQIRSLLSVRMGKEEEKLMIRG
1						LGDIMNNKVFYQHPNLMRALGMHETVMEV
				!		MVNVLGGGESKEITFPKMVANCCRFLCYFCR
			i I			ISRQNQKAMFDHLSYLLENSSVGLASPAMRG
				ļ		STPLDVAAASVMDNNELALALREPDLEKVVR YLAGCGLQSCQMLVSKGYPDIGWNPVEGER
			•			YLDFLRFAVFCNGESVEENANVVVRLLIRRPE
				-		CFGPALRGEGGNGLLAAMEEAIKIAEDPSRD
						GPSPNSGSSKTLDTEEEEDDTIHMGNAIMTFY
						SALIDLLGRCAPEMHLIHAGKGEAIRIRSILRS
}		}]	ļ	j	LIPLGDLVGVISIAFQMPTIAKDGNVVEPDMS
			ľ	ĺ		AGFCPDHKAAMVLFLDRVYGIEVQDFLLHLL
]			}			EVGFLPDLRAAASLDTAALSATDMALALNRY
		į	j			LCTAVLPLLTRCAPLFAGTEHHASLIDSLLHT
i I			ŀ			VYRLSKGCSLTKAQRDSIEVCLLSICGQLRPS
1	1	- 1	Ì			MMQHLLRRLVFDVPLLNEHAKMPLKLLTNH
						YERCWKYYCLPGGWGNFGAASEELHLSRK
		1	ļ			LFWGIFDALSQKKYEQELFKLALPCLSAVAG ALPPDYMESNYVSMMEKQSSMDSEGNFNPQ
		- 1	ŀ			PVDTSNITIPEKLEYFINKYAEHSHDKWSMDK
		i				LANGWIYGEIYSDSSKVQPLMKPYKLLSEKE
	ĺ	1	1	ľ	ľ	KEIYRWPIKESLKTMLARTMRTERTREGDSM
		ĺ			1	ALYNRTRRISQTSQVSVDAAHGYSPRAIDMS
	1			{		NVTLSRDLHAMAEMMAENYHNIWAKKKKM
	ı	i		-		ELESKGGGNHPLLVPYDTLTAKEKAKDREKA
	ļ	ļ	1	J	1	QDILKFLQINGYAVSRGFKDLELDTPSIEKRFA
	1	- 1		1	1	YSFLQQLIRYVDEAHQYILEFDGGSRGKGEHF
		ı				PYEQEIKFFAKVVLPLIDQYFKNHRLYFLSAA
	1		ľ		1	SRPLCSGGHASNKEKEMVTSLFCKLGVLVRH
	Ī				j	RISLFGNDATSIVNCLHILGQTLDARTVMKTG
	ĺ	1			ł	LESVKSALRAFLDNAAEDLEKTMENLKQGQF THTRNQPKGVTQIINYTTVALLPMLSSLFEHI
	-	j	ŀ	Ì	ļ	GQHQFGEDLILEDVQVSCYRILTSLYALGTSK
			ł	i		SIYVERQRSALGECLAAFAGAFPVAFLETHLD
						KHNIYSIYNTKSSRERAALSLPTNVEDVCPNIP
	1	- 1	· ·	- 1	. 1	SLEKLMEEIVELAESGIRYTQMPHVMEVILPM
	Ī	-	-	1		LCSYMSRWWEHGPENNPERAEMCCTALNSE
'	1	- 1	- 1	- 1	1	HMNTLLGNILKITYNNLGIDEGAWMKRLAVF
1		İ	- 1	ļ		SQPIINKVKPQLLKTHFLPLMEKLKKKAATVV
- 1		- 1	1			SEEDHLKAEARGDMSEAELLILDEFTTLARDL
1]	1		İ		YAFYPLLIRFVDYNRAKWLKEPNPEAEELFR
		ı		İ	İ	MVAEVFIYWSKSHNFKREEQNFVVQNEINN
ļ	- 1	- 1		ł		MSFLITDTKSKMSKAAVSDQERKKMKRKGD
1		[-		RYSMQTSLIVAALKRLLPIGLNICAPGDQELIA
-			- 1			LAKNRFSLKDTEDEVRDIIRSNIHLQGKLEDP AIRWOMALYKDLPNRTDDTSDPEKTVERVL
1			- 1]		DIANVLFHLEQKSKRVGRRHYCLVEHPORSK
1	1	J]	1		KAVWHKLLSKQRKRAVVACFRMAPLYNLPR
	-		1	-		HRAVNLFLQGYEKSWIETEEHYFEDKLIEDLA
ļ			1			KPGAEPPEEDEGTKRVDPLHQLILLFSRTALT
						EKCKLEEDFLYMAYADIMAKSCHDEEDDDG

					T 12 4 4 4 4	Amino acid sequence (A=Alanine C=Cysteine,
EQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end nucleotide	D=A spartic Acid. E=Glutamic Acid.
O: of	NO: of	hod	ID NO:	beginning nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
ucl-	peptide		in USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
otide	seq-		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
cq-	uence		914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine,
ence	ł	!	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		1	}	residue of	sequence	V=Tyrosine, X=Unknown, *=Stop codon,
	1	İ	l	peptide	1 - 1	/=possible nucleotide deletion, \=possible
	Į.	ì	1	sequence	<u> </u>	nucleotide insertion
		 	 	Sequence	 	EEEVKSFEEKEMEKQKLLYQQARLHDRGAA
		Ì	1		1	EMVI OTISASKGETGPMVAATLKLGIAILNGG
		ł	1	1		NSTVOOKMLDYLKEKKDVGFFQSLAGLMQS
		1	1			CSVI.DLNAFERONKAEGLGMVTEEGSGEKV
	ļ	ļ	1			LODDEFTCDLFRFLOLLCEGHNSDFQNYLRT
		1	}	ļ	1	OTGNATTVNIIISTVDYLLRVQESISDFYWYY
		1				SGKDVIDEOGORNFSKAIOVAKOVFNTLTEYI
	1	}	1	{	1	OGPCTGNOOSLAHSRLWDAVVGFLHVFAHM
		ì		1		OMKLSODSSOIELLKELMDLQKDMVVMLLS
	i	i	1	1		MLEGNVVNGTIGKOMVDMLVESSNNVEMIL
	1]	j	1		KFFDMFLKLKDLTSSDTFKEYDPDGKGVISK
		1	1			RDFHK AMESHKHYTOSETEFLLSCAETDENE
	ł	1	1	,		TLDYEEFVKRFHEPAKDIGFNVAVLLTNLSEH
	1	1	}	}	1	MPNDTRLOTFLELAESVLNYFQPFLGRIEIMG
	1 '	{	-	1	Ì	SAKRIERVYFEISESSRTQWEKPQVKESKRQFI
		1	1		1	FDVVNEGGEKEKMELFVNFCEDTIFEMQLAA
	1		1		1	OISESDLNERSANKEESEKERPEEQGPRMAFF
	1	1	1	ì	1	SILTVRSALFALRYNILTLMRMLSLKSLKKQM
	ł	١		ì	ļ	KKVKKMTVKDMVTAFFSSYWSIFMTLLHFV
	1	ſ			ļ	ASVERGEFRIICSLLLGGSLVEGAKKIKVAELL
	1	1			}	ANMPDETODEVRGDGEEGERKPLEAALPSED
				1	1	1.TDLKELTEESDLLSDIFGLDLKREGGQYKLIP
	1	1		1	1	HNPNAGLSDLMSNPVPMPEVQEKFQEQKAK
		1	ļ			EEEKEEKEETKSEPEKAEGEDGEKEEKAKED
		1		1	į.	KGKQKLRQLHTHRYGEPEVPESAFWKKIIAY
		ļ				OOKLLNYFARNFYNMRMLALFVAFAINFILL
	1	1	1	1		FYKVSTSSVVEGKELPTRSSSENAKVTSLDSS
	ł	ì	i			CHRILAVHYVLEESSGYMEPTVRILPILHTVISF
					ì	FCUGVVCLKVPLVIFKREKEVARKLEFDGLYI
		1	1			TEOPSEDDIKGOWDRLVINTOSFPNNYWDKF
		1		j	1	VKRKVMDKYGEFYGRDRISELLGMDKAALD
		-	- [1 '		FSDAREKKKPKKDSSLSAVLNSIDVKYQMW
	}	•	1	İ	i	KLGVVFTDNSFLYLAWYMT
			1		767	LPELNGRGAGLRRAEPSERGGGAERTQQVAA
553	1903	A	4199	31	/6/	I PL SHGHSHGGGGCRCAAER/VGAARGSAAC
	.]	- 1	1	1	ļ.	AYGLYLRIDKGRLQCLNESREGSGRGVFKPW
İ		ì		1		ERAD\DRSKFVESDADEELLFNIPFTG\HVKLK
}	}	1	1		-	GIIMGEDDDSHPSEMRLYKNIPQMSFDDTER
1	1	1				EPPOTESI NRDLTGELEYATKISRFSNVYHLSI
l		1		1	ļ	HISKNFGADTTKVFYIGLRGEWTELRRHEVTI
}	1	}	-	1		CNYEASANPADHRVHQVTPQTHFIS
1	1					GIPCTEMGNFDNANVTGEIEFAIHYCFKTHSL
554	1904	A	4200	1	961	EICIKACKNLAYGEEKKKKCNPYVKTYLLPD
	1			. 1	1	RSSQGKRKTGVQRNTVDPTFQETLKYQVAPA
Į.	- 1	1	1	1		QLVTRQLQVSVWHLGTLARRVFLGEVIIPLAT
ļ	}	- 1	l	ŀ		WDFEDSTTQSFRWHPLRAKADKYEDSVPQS
		l l	1	1		NGELTVRAKLVLPSRTRKLQEAQEGTDQPSL
1	- 1	ļ	1	l l		HGQLCLVVLGAKNLPVRPDGTLNSFVKGCLT
1	1	1	}	}	- 1	LPDQQKLRLKSPVLRKQACPQWKHSFVFSGV
1	1	l	1	1		TPAQLRQSSLELTVWDQALFGMNDRLLGGTN
1		1		1	į.	RLGSKGDTAVGGDACSQSKLQWQKVLSSPN
1	ļ	- 1	1	1	1	
1	i	1	- 1	l		LWTDMTLVLH
555	1905		4211	331	2419	KENKKARNLRMNQSRSRSDGGSEETLPQDH
725	1,703	'`		1		NHHENERR WQQERLHREEAYYQFINELNDE
1		1	1	1		DYRLMRDHNLLGTPGEITSEELQQRLDGVKE
	1	ł		1		QLASQPDLRDGTNYRDSEVPRESSHEDSLLE
1	1	1	- 1	ì	- 1	WLNTFRRTGNATRSGQNGNQTWRAVSRTNP
1	- 1				1	Language por premiural Description (FI)
	1	}	1		1	NNGEFRESLEIHVNHENRGFEIHGEDYTDIPLS
			}	-		DSNRDHTANRQQRST\SPVARRTRSQTSVNFN GSSSNIPRTRLASRGQNPAEGSFSTLGRLRNG

SEQ III NO: of nucl- eotide seq- uence		Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GGAAGIPRANASRTNFSSHTNQSGGSELRQRE GQRFGAAHVWENGARSNVTVRNTNQRLEPI RLRSTSNSRSRSPIQRQSGTVYHNSQRESRPV QQTTRRSVRRRGRTRVFLEQDRERERRGTAY
556	1906	A	4212	3	462	TPFSNSRLVSRITVFEGEESSRSSTAVRRHPTIT LDLQVR/RIRPGENRDRDSIANRTRSRVGLAE NTVTIESNSGGFRRTISRLERSGIRTYVSTITVP LRRISENELVEPSSVALRSILRQIMTGFGELSSL MEADSESELQRNGQHLPDMHSELSNLGTDN NRSQHREGSSQDRQAQGDSTEMHGENETTQP HTRNSDSRGGRQLRNPNNLVETGTLPILRLAH FFLLNESDDDDRIRGLTKEQUDNLSTRHYEHN SIDSELGKICSVCISDYVTGNKLRQLPCMHEF HIHCIDRWLSENCTCPICRQPVLGSNIANNG
					462	LQRQRQHPAAAPAVPVRCFTFCFTDIVIMPKR KSPENTEGKDGSKVTKQEPTRRSARLSAKPA PPKPEPKPRKTSAKKEPGAKISRGAKGKKEEK QEAGKEGTAPSENGETKAEEIHISRSTVNVST SRGTPPSTLSVKGQIETVRVKGTEN
557	1907	A	4213	774	507	ARRFSCLTLQTSWGHRH\GPPRP\ANFVFLVET GFLHIGQAGHKLPTSGDPPASASQSARITGMS HRTWFLASFLIDSCKNFIVYKIMYTL
558	1908	A	4225	3	1253	TYRHAEREHPETSSATKVSYDYRHKRPKLLD GDQDFSDGRTQKYCKEEDRKYSFQKGPLNRE LDCFNTGRGRETQDGQVKEPFKPSKKDSIAC TYSNKNDVDLRSSNDKWKEKKKKEGDCRKE SNSSSNQLDKSQKLPDVKPSPINLRKKSLTVK VDVKKTVDTFRVASSYSTERQMSHDLVAVG RKSENFHPVFEHLDSTQNTENKPTGEFAQEIIT IIHQVKANYFPSPGITLHERFSVKMADIHKADV NEIPLNSDPEIHRRIDMSLAELQSKQAVIYESE QTLIKIIDPNDLRHDIERRKERLQNEDEHIFHI ASAAERDDQNSSFSKNYTTQRKDIITHKPFEV EGNHRNTRVRPFKSNFRGGRCQPNYKSGLVQ KSLYIQAKYQRLRFTGPRGFITHKFRERLMRK KKVP
559	1909	A	4235	1	323	KFSIPFFLRWSFTLVVPRLEGNDMISVHCNLGL LGLSHSPASASQVGGITGTQHHTGLIFGFLIET EFHHVGQAGLELLTSGDPPALAFQSAGITGVS HHAWLQVLNS
560	1910	A -	4246	2	1569	TLSLLERVLMKDIVTPVPQEEVKTVIRKCLEQ AALVNYSRLSEYAKIEGKKREMYELPVFCLA SQVMDLTIQNQKDAENVGRLITPAKKLEDTIR LAELVIEVLQQNEEHHAEAFAWWSDLMVEH AETFLSLFAVDMDAALEVQPDTWDSFPLFQ LL'NDFLRTGLLICGNGKVFHKHLQDLFAPLVV RYMWDLDGSSPIAQSIHRGLLSRESWEPVNN GSGTSEDLFWKLDALQTFIRDLIIWPEEEFGK HLEQRLKLMASDMIESCVKRTRIAFEVKLQK TSSIQQIFRVPQFNMAPCFNVMGLMAKGSIQP KL'CSMEMGQEFAKMWHQYHSKIDELIEETV KEMITLLVAKFVTILEGVLAKLSRYDEGTLFS SFLSFTVKAASKYVDVPKPGMDVADAYVTF VRHSQDVLRDKVNEEMYIERLFDQWYNSSM NVICTWLTDRMDLQLHIYQLKTLIRMVKKTY RDFRLQGVLDSTLNSKTYETIRNRLTVEEATA SVSEGGGLQGISMKDSDEEDEEDD
561	1911	A	4257	1300	654	SELVQFLLIKDQKKIPIKRADILKHVIGDYKDI FPDLFKRAAERLQYVFGYKLVELEPKSNTYIL

SEQ ID	SEQ ID	Mat	SEO	Dendistad	I Davidiana and	I Amino anid announce (AmAlusius C. C. a.)
NO: of	NO: of	Met	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	поц	in NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
cotide		1	1		location	F=Phenylalanine, G=Glycine, H=Histidine,
	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uaice	}	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		į	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ļ		-	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ì	1	İ	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ŀ			peptide		/-possible nucleotide deletion, \-possible
 	<u> </u>		 	sequence	ļ	nucleotide insertion
		ł	1	ŀ	1	INTLEPVEEDAEMRGDQGTPTTGLLMIVLGLI
		1]	İ	FMKGNTIKETEAWDFLLAL\GVYPTKKHLIFG
					ļ	DPKKLITEDFVRQRYLEYRRIPHTDPVDYEFQ
		1				WGPRTNLETSKMKVLKFVAKVHNQDPKDW
					<u> </u>	PAQYCEALADEENRARPQPSGPAPSS
562	1912	A	4260	1	1498	MVTWLYRFLPTSNMAAKLRSLLPPDLRLQF
	Ì		Ī			WLHARLQKCFLSRGCGSYCAGAKASPLPGK
			1	1	1	MAMGLMCGRRELLRLLQSGRRVHSVAGPSQ
		1			ł	WLGKPLTTRLLFPAAPCCCRPHYLFLAASGPR
	1	ĺ			İ	SLSTSAISFAEVQVQAPPVVAATPSPTAVPEV
1		ľ				ASGETADVVQTAAEQSFAELGLGSYTPVGLI
1				1		QNLLEFMHVDLGLPWWGAIAACTVFARCLIF
	}	1		}	{	PLIVTGQREAARIHNHLPEIQKFSSRIREAKLA
	1	1				GDHIEYYKASSEMALYQKKHGIKLYKPLILPV
	1	l	ļ		1	TQAPIFISFFIALREMANLPVPSLQTGGLWWF
1		j			j	QDLTVSDPIYILPLAVTATMWAVLELGAETG
	İ]			VQSSDLQWMRNVIRMMPLITLPITMHFPTAV
].						FMYWLSSNLFSLVQVSCLRIPAVRTVLKIPQR
		ļ	}			VVHDLDKLPPREGFLESFKKGWKNAEMTRO
	ĺ	1				LREREQRMRNQLELAARGPLRQTFTHNPLLQ
Ì	i		1 1			PGKDNPPNIPSS\SSSSSKPKSKYPWHDTLG
563	1913	Α	4265	623	116	MGGLAPTQTLEPT\REYQNTQLSVSYLLPEQN
	""			023	***	THGTRRTLSSGPSNNLPLPLSSSATMPSMOCK
			i i			HRSPNGGLFRQSPVK/TPPIPMSFQPVPGGVL
		ĺ				PRGSGNPPHGTSILTAPPALLPHPPTHPTQQSF
·			1		1	
				,		LIQENNNTNHTHSHTHTYTETLSFFLYICVNN DRMEWGKSVF
564	1914	A	4270	3 .	368	
501	1777	l '`	12/0	٠, •	300	ILKRKLSSLNSEVSTIQNTRMLAFKATAQLFIL
		l				GCTWCLGLLQVGPAAQVMAYLFTINSLQGF
						FIFLVYCLLS\QQVQKQYQKWFREIVKSKSES
565	1915	A	4288	83	406	ETYTLSSKMGPDSKPSEGDVFPRTSE
202	1913	Α	4200	83	406	RNSRPLWCSPPASQPRQAPVSQSCCCPLPSSSS
			1			PPSALLAPTKPRALGTLRLYECSPELGTTMLP
					i	PAWLLMLCQAPRPQDPDPRLTQPEKSLQEAP
500	1016		1000			GQTGASRTPRT
566	1916	Α	4298	1041	229	LNSSQKLACLIGVEGGHSLDSSLSVLRSFYVL
						GVRYLTLTFTCSTPWAESSTKFRHHMYTNVS
			1 [GLTSFGEKVVEELNRLGMMIDLSYASDTLIRR
				i		VLEVSQAPVIFSHSAARAVCDNLLNVPDDILQ
			1			LLKKNGGIVMVTLSMGVLQCNLLANVSTVA
						DHFDHIRAVIGSEFIGIGGNYDGTGRFPQGL\E
]]	l		DVSTYPVLIEELLSRSWSEEELQGVLRGNLLR
						VFRQVEKVREESRAQSPVEAEFPYGQLSTSCH
						FHLGASEWTPRLLIWR
567	1917	A	4299	1	1106	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE
		1	i			DFPETSEPVWILGRKYSIFTEKDEILSDVASRL
				1	ļ	WFTYRKNFPAIGGTGPTSDTGWGCMLRCGO
				1	ļ	MIFAQALVCRHLGRDWRWTQRKRQPDSYFS
			j	ļ		VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGQ
			}	1	l	WYGPNTVAOVLKKLAVFDTWSSLAVHIAMD
				ſ	ĺ	NTVVMEEIRRLCRTSVPCAGATAFPADSDRH
	•			ļ		
					ļ	CNGFPAGAEVTNRPSPWRPLVLLIPLRLGLTD
			j			INEAYVETLKHCFMMPQSLGVIGGKPNSAHY
ļ] ']		ļ	FIGYVGEELIYLDPHTTQPAVEPTDGCFIPDES
						FHCQHPPCRMSIAELDPSIAVVRGGHLSTQAF
560	1010	_	4200	2012	1040	GAECCLGMTRKTFGFLRFFFSMLG
568	1918	A	4300	2012	1843	SRKFLTITPIVLYFLTSFYTKYDQIHFVLNTVS
	1010		1205	486		LMSVLIPKLPQLHGVRIFGINKY
569	1919	A	4302	186	531	WTFCLFL/WWVPESARWLLTQGHVKEAHRY

SEQ ID	SEQ ID	Met	SEQ	D. dias	18 0 t	
NO: of	NO: of	hod	ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
		i i	}	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	Ì	,	peptide	•	/=possible nucleotide deletion, \=possible
		<u> </u>	<u> </u>	sequence		nucleotide insertion
1						LLHCARLNGRPVCEDSFSQEVRVNVCVSMHI
1					1	CVWWGVGCVKCLPPRAHHIWQEKPLGPHRT
					<u> </u>	VTESKLEAEGKTKEKAREKERKKKS
570	1920	Α	4308	3	869	RSGQGKVYGLIGRRRFQQMDVLEGLNLLITIS
}	ļ	J	J]	ļ	GKRNKLRVYYLSWLRNKILHNDPEVEKKQG
		ļ				WTTVGDMEGCGHYRVVKYERIKFLVIALKSS
		1	}			VEVYAWAPKPYHKFMAFKSFADLPHRPLLV
						DLTVEEGQRLKVIYGSSAGFHAVDVDSGNSY
		l	1			DIYIPVHIQSQITPHAIIFLPNTDGMEMLLCYE
		1				DEGVYVNTYGRIIKDVVLQWGEMPTSVAYIC
1		l				SNQIMGWGEKAIEIRSVETGHLDGVFMHKRA
	ĺ	ĺ				QRLKFLCERNDKVFFASVRSGGSSQVYFMTL NRNCIMNW
571	1921	A	4309	9	524	ASREMDVTKVCGEMRYQLNKTNMEKDEAE
	1 -7		1507		324	KEHREFRAKTNRDLEIKDQEIEKLRIELDESK
	i	ļ				QHLEQEQQKAALAREECLRLTELLGESEHOL
İ						HLTRQEKDSIQQSFSKEAKAQALQAQQREQE
		j				LTQKIQQMEAQHDKTENEQYLLLTSQNTFLT
1	}		ł			KLKEECCTLAKKLEOISO
572	1922	Α	4318	1	1119	GATPLGSVGGRTGKMDAATLTYDTLRFAEFE
ı						DFPETSEPVWILGRKYSIFTEKDEILSDVASRL
	1					WFTYRKNFPAIGGTGPTSDTGWGCMLRCGQ
1	ļ	ļ				MIFAQALVCRHLGRDWRWTQRKRQPDSYFS
		}				VLNAFIDRKDSYYSIHQIAQMGVGEGKSIGO
			ŀ			WYGPNTVAQVLKKLAVFDTWSSLAVHIAMD
						NTVVMEEIRRLCRTSVPCAGATAFPADSDRH
		1	}			CNGFPAGAEVTNRPSPWRPLVLLIPLRLGL\T
i						DINEAYV\ETL\KHCFHGWPQFPG/VVHREGK
			l [PNSAHYFIGYVGEELIYLDPHTTQPAVEPTDG
						CFIPDESFHCQHPPCRMSIAELDPSIAVVRGGH
573	1923	A	4333	363	1066	LSTQAFGAECCLGMTRKTFGFLRFFFSMLG
3/3	1923	А	4333	303	1000	GGVPVGLASKPFQILYGHTNEVLSVGISTELD
						MAVSGSRDGTVIIHTIQKGQYMRTLRPPCESS
1						LFLTIPNLAISWEGHIVVYSSTEEKTTLK\ERM
			1		-	HYICFSINGKYLGSQILKEQVSDICIIGEHIVTG SIQGFLSIRDLHSLNLSINPLAMRLPIHCVCVT
1		1			ļ	KEYSHILVGLEDGKLIVVGVGKPAEVKPSISN
		į			1	FISHAVGDYFGSPSFQLIEKSPLGINKLKAKFD
<u> </u>						FSKGSK
574	1924	Α	4346	359	1234	MDTLEEVTWANGSTALPPPLAPNISVPHRCLL
	i	i		l	l	LLYEDIGTSRVRYWDLLLLIPNVLFLIFLLWK
				ĺ		LPSARAKIRITSSPIFITFYILVFVVALVGIARA
1				l		VVSMTVSTSNAATVADKILWEITRFFLLAIEL
∤						SVIILGLAFGHLESKSSIKRVLAITTVLSLAYSV
				Į.		TQGTLEILYPDAHLSAEDFNIYGHGGROFWL
				l		VSSCFFFLVYSLVVILPKTPLKERISLPSRRSFY
1	ļ			1	1	VYAGILALLNLLQGLGSVLLCFDIIEGLCCVD
575	1005	<u> </u>	10.55			ATTFLYFSFFAPLIYVAFLRGFFGSEPKILF
575	1925	A	4360	2038	1512	GCWWRHPWLASQRDCLDCRIQLAEKFVKAV
				1	ĺ	SKPSRPDMNPIRVKEVYRLEEMEKIFVRLEM
]]	Į			ţ	l	KIIKGSSGTPKLSYTGRDDRHFVPMGLYIVRT
	1			- 1	Į	VNEPWTMGFSKSFKKKFFYNKKTKDSTFDLP
	•			j	Į	ADSIAPFHICYYGRLFWEWGDGIRVHDSQKP
576	1926	A	4365	69	500	QDQDKLSKEDVLSFIQMHRA
ا "'ا	1920	^	4303	לס	500	QVEGRQGREVKRTAWRISPVWRPARCRRRST
			-		i	PQP/PE/PGAQQQERHRQGEAPMQALDPRAEP
[J		1			GPQAQSHAACQPEPEPPRVLLDPTAARGGVQ
1 1	- 1	l	į	1	1	GRP/GLSRHPGLAPHPQTHTPWPQSGRLPCAS
	!					EPLPLGGIRPTPGLEPKGRDLM

SEQ ID	SEQ ID	Met	SEO	D 3!-4-3	l b - 3: 4 - 1 - 4	14-1
NO: of	NO: of	hod	ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide	"""	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	l	ŀ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	ĺ			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
İ			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
		ł	1	peptide	1	/=possible nucleotide deletion, \=possible
]		sequence		nucleotide insertion
577	1927	Α	4366	785	502	SAPPKKKNGVLFLSPRLKSSGAIWVHSTPTLW
		Ì				ASSNSRASTPKVAGITGARPHARIIFVFLIEMG
			<u> </u>			FHNVGQAGL/DTLTLVICPPQPPKLLGLQM
578	1928	A	4367	1	221	FFFFLKKSRCVTQAGVQG\PISLHPPPPGFKRF
						SRLSLLSSWDYRHP/HAANFCIFSRDG\VSPYW
	4000	<u> </u>	ļ			SGWSRTPDLR
579	1929	A	4383	1	224	FETESHSVTQAGMQWHNLGSLQPMP/PGLKR
		([FSCLRLQSSWDHRHAPPHLAHFCIFSRDGVSP
600	1000	ļ.,	1222			CWPGWSSTPDLK
580	1930	A	4397	410	94	SRLKPYSTNVTAKKLPATNIPNLDCFTAKLYQ
	l				1	\VFKKGNHILHELFQNKEEGAFPNS/FYEASFT
					f	LRPKSDRDIAKEESYSTISLLSTDTKILMSKYK
581	1931	A	4414	670		QLKSSDL
J01	1931	^	4414	670	3	VLVHRQCGGILRLRRKEAVSVLDSADIEVTDS
			1			RLPHATIVDHRPQHRWLETCNAPPQLIQGKA
			ł			RSAPKPSQASGHFSVELVRGYAGFGLTLGGG
						RDVAGDTPLAVRGLLKDGP\AQRCGRLEVGD LVLHINGESTQGLT\HAQAVERIRAGGPQLHL
						VIRRPLETHPGKPRGVGEPRKGVVPSWPDRSP
						DPGGPEVTGSRSSSTSLVQHPPSRTTLKKTRG
						SPE
582	1932	Α	4424	194	449	VLYIRKKKRLEKLRHQLMPMYNFDPTEEQDE
	-70-	••		124	777	LEQELLEHGRDAASVQAATSVQAMQGKTTL
				· i		PS\QGPLQRPSRLVFT\DVANAIHV
583	1933 .	A	4435	1	166	APGPPVPPPGSPPEQMPGPCPASMPP/DPPPGS
				_	100	PPEQMPGPCPVSAPP/GPPPGSPPEQMPGPCPV
				,		SAPPALLODTSV
584	1934	Α	4439	1	628	SATPQQPSAPQHQGTLNQPPVPGMDESMSYQ
				•		APPQQLPSAQPPQPSNPPHGAHTLNSGPQPGT
						APATQHSQAGPATGQAYGPHTYTEPAKPKK
]			GQQLWNRMKPAPGT\EVSSSTSRSDPLLLPPR
			1			ALAPTQRASTVVLAPSPT/SEKVQNHSGSSAR
						GNLSGKPDDWP/LGHERVCGALLHRL*VGGG
						QGPHGKAAQGGAAGAAGRLGLYH
585	1935	Α	4463	10	144	HKPVTNSRDTQEVPLEKAKQVLKIIATFKHTT
						SIFDDFAHYEKRQ
586	1936	Α	4464	1309	103	LNAESYVSFTTKLDIPTAAKYEYGVPLQTSDS
			l i			FLRFPSSLTSSLCTDNNPAAFLVNQAVKCTRK
- 1						INLEQCEEIEALSMAFYSSPEILRVPDSRKKVPI
						TVQSIVIQSLNKTLTRREDTDVLQPTLVNAGH
			· 1		j	FSLCVNVVLEVKYSLTYTDAGEVTKADLSFV
ļ				ļ	İ	LGTVSSVVVPLQQKFEIHFLQENTQPVPLSGN
	[ſ	ľ	PGYVVGLPLAAGFQPHKGSGIIQTTNRYGQLT
· 1	•	. 1	i i	-		ILHSTTEQDCLALEGVRTPVLFGYTMQSGCK
}		Ť				LRLTGALPCQLVAQKVKSLLWGQGFPDYVA
İ						PFGNSQGP/ADMLDWVPIHFITQSFNRKDSCQ
- 1						LPGALVIEVKWTKYGSLLNPQAKIVNVTANLI
j	1			1		SSSFPEANSGNERTILISTAVTFVDVSAPAEAG
587	1937	<u></u>	4471	614	387	FRAPPAINARLPFNFFFFFV
١	-///	Λ	77/1	014	301	LLGRASAC/LQLQSSW/D/HRPMLPYLANFVF
}	į			ļ	ļ	CKDR/SFTWLPRLVLNSWLQVILLPWPPTGCD
588	1938	Ā	4480	1720	1458	NKHEPPCPATKRRHSGSI
200	1,700	^	1400	1/20	1458	HDLGSLQPPPPGFKRFSCLSLPSSWDYRLMPP
Ì	Ì	l		ļ		CPANFCIIII/DFLVETGFHHVGQASHELLTSGD
589	1939	A	4487	922	222	PPTSASQSAGITGMSYHTWFGES
507	1737	^	740/	722	332	APVTTSPRVGQPW/RTALALRSLYRARPSLRC
				1	1	PPVELPWAPRRGHRLSPADDELYQRTRISLLQ
	ŀ			j	1	REAAQAMYIDSYNSRGFMINGNRVLGPCALL PHSVVONDNGSHODITEDSESI EVI I EDDIEI
						PHSVVQWNVGSHQDITEDSFSLFWLLEPRIEI

NO. of No. of nod not	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Decide Seq							
Sequence							
Page Uenice			1				
gence			1				M-Methionine No Assessing D-D-11-
amino acid residue of peptide residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence per seque		100.00	Ι.				O-Clutomine D-Aminine C-Comine,
residue of peptide pep	delice		'	714			Q=Glutamme, K=Arginine, S=Serine,
Peptide Pept	1		l	ł			
	1	İ		ļ		sequence	Y=1 yrosine, X=Unknown, *=Stop codon,
1940 A 4492 1 472	1	l		i		}	
DTPNACATIFIC LEGRVTGAAL DTPNACATIFIC STALOQAQ		<u> </u>			sequence		
1940 A 4492 1 472	1		1	į			VVVGTGDRTERLQSQVLQAMRQRGIAVEVQ
1940 A 4492 1 472			1	İ			DTPNACATFNFLCHEGRVTGAALIPPPGGTSL
PPSCLSLPSSWDYRRPL RPANFFVT VETGIP RPSRDGLDL17/S/GDPTSASQSAGTITOVIST RPSRDGLDL17/S/GDPTSASQSAGTITOVIST RPSRDGLDL17/S/GDPTSASQSAGTITOVIST RPSRDGLDL17/S/GDPTSASQSAGTITOVIST VTHQGGLPGFRYAFSASGQREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGLPGFRYAFSAGGREAS/GDPGR THQGGGLSLANAWGLSVGAKGKWKPLEFL CTLAATHMLNVAVPIATSVVQLRQRDFD EWREGLCKVPVSTSTLATSSYSH MWMVCWPVYRLISAAKQAGHTVMGIWM GSFLSALPAVGWHDTSSRTYTHGCFRYAFS GGGGVCFLLVGGSVAMGVICTAIALFQTL AVQVGQADHRAFTVTIVVEDAGGRACVM GFSLSALPAVGWHDTSSRTYTHGCFRYAFS THQGGGGACVM GFSLSALPAVGWHDTSSRTYTHGCFRYAFS GFSLSALPAVGWHDTSSRTYTHGCFRYAFS GFSLSALPAVGWHDTSSRTYTHGCFRYAFS GFSLSALPAVGWHDTSSRTYTHGCFRYAFS GFSLSALPAVGWHDTSSRTYTHGCFRYAFS GGYKWVGLWLIAHPPRWGLAGADGACVM GGYKWVGLWLIAHPPRWGLAGADGACVM GGYKWVGLWLIAHPPRWGLAGADGACVM GGYKWVGLWLIAHPPRWGLAGADGACVM GGYKWVGLWLIAHPPRWGLAGADGACVM THGGGR THGGG		12.12				<u> </u>	
	590	1940	Α	4492	1	472	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT
SPI	l	1]				PFSCLSLPSSWDYRRPPLRPANFFVFLVETGFP
1941 A 4495 1444 1116		1	Í		1	[RFSRDGLDLLT/S/GDPPTSASQSAGITGVSHR
1941 A 4495 1444 1116	l				}		ARPKRIGEPRRKCGNAVVWPSTSLGDHRVTS
1941				1			VPHQGGLPGPIRVAPSSAGOREASOGPPGR
MEYYADTERNEIMSFAGTWVELEAIT, SKLM	591	1941	A	4495	1444	1116	
LKDNWVBDTIPQGAVPCTATAEGMKRLLFAL	İ						
S92	ľ		i	1			
1942 A		l		l		•	
VCGGLSLLANAWGILSVGAKQKKWKPLEFI LCTLAATHMLNVAVPIATYSVVQLRQRPDFI EWNEGLCKVPVSTYTLIATCTSSYYSLSYHR MWMVCWPVNYRLSNAKKQAGHTVMGIWM GSFILSALPAVGWHDTSETYTIGCRPIVAEI GLGFGVCFLLLVGGSVAMGVICTAALFOTL AVQVGRQADHRAFTVPTIVVEDAQGKRRSI DGSEPAKTSLQTTGL-VTIVPTIPOLAMGFPVL GFFSLADTHLSDLPYTWGDRDSGGACVM GFFSLADTHLSDLPYTWGDRDSGGACVM FFFEAESCSVPQAGVQRPDLGWLHAPPPGSC HFPASASQVAGTTHARHHTQLIPAFLVENGL C	592	1942	A	4496	2	919	
LCTLAATHMILNVAVPIATYSVQLIRQRPDF	}	., .	1	1	~	7.7	
EWNEGLCK VFVSITYITLITATCS YSLSYHR				1			I CTL A ATEM ANY AVENTATIVE OF PROPERTY
	1						EUNECI CVVEVETEVET EL ATOCCUTOL CUATR
GSFILSALPAVGWHOTSERFYTHGCRIVAGE							EWNEGLCKYFYSIF I LILAICFSVI SLSYHK
	1			!			MWMVCWPVNYKLSNAKKQAGHIVMGIWM
S93 1943 A 4506 2 193 FFFEAESCSVPQAGVQRPDLGWLHAPPPVSGC HFPASASQVAGTTHARHTQLIFVAFUENDSGACVM HPASASQVAGTTHARHTQLIFVAFUENDSG ESTEIPPYVMKCPSPSGGACVM GRYPASASQVAGTTHARHTQLIFVAFUENDSG CONTROL OF CONTROL O							GSFILSALPAVGWHD1SERFY1HGCRFIVAEI
593 1943 A 4506 2 193 FFFEAESCSVPQAGVQRPDLGWLHAPPPAGSC HFFAESASQVAGTTHARHHTQLIFVAFLVENGL C C KMAGGVRPLGRULGWLHAPPPAGSC HFFAESASQVAGTTHARHHTQLIFVAFLVENGL C C KMAGGVRPLGRULGWLHAPPPAGSC HFFAESASQVAGTTHARHHTQLIFVAFLVENGL C C KMAGGVRPLRGLRALCRVLLFLSQFCILSGG ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNFS CTYGRPVTFDCAVKPSVTCVDQDFKSQKNFII NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ RYPANCTVRIDHVHCLGNSTFFKMLYCNWT GGYKWVGLWLLHIPRWGLGADRPYLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI 595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHILLGSNN SPASASRV/AGNIGARHHTQQIFVLLVQMRVH YVGQDGLDLL/NLMHIPPRSFVLGLQA FFFKMESYSVARLECSGAISAPCNLHILLGSNN SPASASRV/AGNIGARHHTQQIFVLLVQMRVH YVGQDGLDLL/NLMHIPPRSFVLGLQA HASDFIL/PYFFLVALLAKQRQFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NILMELLVQKKKQLEAESIHAAQLQILMEFILK VARNKREQLEQIGKELSVLEEDIKRVEMS GLYSPVSEDSTVPQFEAPSPSHSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECUS KFATRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKRVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHERCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASEAKANVCCVKRSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHKKAVSYAKFVSG EEIVSASTDSQLKLWNVCKPYCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDLDRKKEDTNEFVSAV CWRALPDGESNVLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNILAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIAANSQGTIKVLELV CWRALPDGESNNLIHLIGSSDPFTS ASQVAEITSVRHITTWLIFCCLGQMGFHHVYGE QAGLELTSWDPAIDFQSAGIIGMSPHAWPP							GLGFGVCFLLLVGGSVAMGVICTAIALFQTL
1943 A 4506 2 193	1 .		İ				
1943 A 4506 2 193 FFFEAESCSVPQAGVQRPDLGWLHAPPPGSC HFPASASQVAGTTHARHHTQLIFVAFIQL C HFPASASQVAGTTHARHHTQLIFVAFIQL C C HFPASASQVAGTTHARHHTQLIFVAFIQL C C KMAGGVPLRGLRALCRVLLFLSQFCILSGG ESTEIPPYVMKCPSNGLCSRLPADCIDCTINFS CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFII NMTCRFCWQLPETDVECTINSTSCMTVSCPRQ RYPANCTVRIDHVHCLGNRTFPKMLYCNWT GGYKWYYGLWLIRHHPRWGLGADRFYYLOP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI 595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRVAGNIGARHHTQQIFVLLVQMRVH YVQQDGLDLLNLMIHPPRSPKVLDQA 596 1946 A 4513 3 1674 HASDHLYPNELILKQKQRFEEKBFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLLS KFNTRYNSVRPLATLSYASDLYNGSQVKSLV FEFDRCDYFALGVIKKKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHERKCWSV DFNLMDFKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKWNVGFWYCLRSFKGHIN EKNFVGLASNGGTVACSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTINEFVSAV CWRALPDGESDVLANNSQGTVKVLEVV CWRALPDGESDVLANNSQGTKVLEEV CWRALPDGESDVLANNSQGTKVLEEV CWRALPDGESDVLANNSQGTKVLEEV CWRALPDGESDVLANNSQGTKVLEEV CWRALPDGESDVLANNSQGTKVLEEV ASQVAEITSVRHITWLICTSCDGMGFHHVGE QAGLELLTSWDPALLPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAAGVFRCCRLPGAYFFSF]				•		
1944 A 4507 1327. 647 KMAGGVRPLRGLRALCRVLLFLSQFCILSGG ESTEIPPYVMKCPSNGLCSRLPADCIDCTINFS CTYGRPVTFDCAVKPSVTCVDQDFKSQKNFII NMTCRFCWQLPFITDYECTNSTSCMTVSCPRQ RYPANCTVRDIVHCLGNRTFPKMLYCNWT GGYKWVYGLWLLRHIPPRWGLGADRFYYLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLY1 595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRV/AGNIGARHITQQIFVLLVQMRVH YVQQDGLDLLNLMIHPPRSPKVLGLQA HASDHLYPNFLVNEILLKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDKRVEEMS GLYSPVSEDSTVPQFEAPSPHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KFITRYNSVRPLATLSTYSADLYNGSQVKSLV FEFDRCDYFALAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHERCKWSV DFNLMDPKLLASGDDAKVKL WSTNLDNSV ASIEAKANVCCVKFSPSSRYH.AFGCADHCV HYYDLRNTKQPIMVFKGHKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPVCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDREDDTNEFVSAV CWRALPDGESNVLLANNSQGTVKLLV CWRALPDGESNVLLANNSQGTNKVLELV CWRALPDGESNVLLANNSQGTNKVLELV GWRALPDGESNVLLANNSQGTNKVLELV ASQVAEITSVRHITWLIFCLLGQMGFHHVGE QAGLELLTSWDPALLPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAAGVFRCRLPGAYFFSF							
C C KMAGGVRPLRGLRALCRVLLFLSQFCILSGG ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNFS CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFII NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ RVPANCTVRQIDHVHCLGRNTFPKMLYCNWT GGYKWVYGLWLLRHHPRWGLGADRFVYLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI	593	1943	Α	4506	2	193	
1944							HFPASASQVAGTTHARHHTQLIF\AFLVENGL
ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNFS CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFII NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ RYPANCTVRDNVHCLGNRTFFKMLYCNWT GGYKWVYGLWLLRHIPRWGLGADRFYYLGP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYYG PADGSLVI 595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRVIAGNIGARHITQQIFVLLVQMRVH YVGQDGLDLLNIMHPPRSPKVLGLQA 596 1946 A 4513 3 1674 HASDHLYPNFLVNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEASHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWVTNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECUS KFITRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVVEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVLWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDVIAGCSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPPGESNVLIAANSQGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHILGSSDPFTS ASQVAEITSVRHITWLIFCRLGQMGFHHVGE QAGLELLTSWDPAILPSQASGIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFRCRLPGAYFFSF							
ESTEIPPYVMKCPSNGLCSRLPADCIDCTTNFS CTYGKPVTFDCAVKPSVTCVDQDFKSQKNFII NMTCRFCWQLPETDYECTNSTSCMTVSCPRQ RYPANCTVR\DHVHCLGNRTFPKMLYCNWT GGYKWVYGLWLLRHIPRWGLGADRF\YLGP VAGTASGKLFSFGLGIWTLIDVLLIGVGYVG PADGSLYI 595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRV\AGNIGARHHTQQIFVLLVQMRVH YVGQDGLDLLNLMHPPRSPKVLGLQA YVGQDGLDLLNLMHPPRSPKVLGLQA HASDHLYPNFLVNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTMSRISDDSRTASQLDEFQECLS KPITRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDYIAGOSENNSLYLYYKGLS STLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPPGESNVLLAANSQGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHILLGSDPFTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQASGIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	594	1944	Α	4507	1327.	647	KMAGGVRPLRGLRALCRVLLFLSOFCILSGG
NMTCRFCWQLPETDYECTINSTSCMTVSCPRQ RYPANCTVRIDHVHCLGNRTFPKMLYCNWT GGYKWVYGLWLLRHIPRWGLGADRPYLOP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI	\$						
NMTCRFCWQLPETDYECTINSTSCMTVSCPRQ RYPANCTVRIDHVHCLGNRTFPKMLYCNWT GGYKWVYGLWLLRHIPRWGLGADRPYLOP VAGTASGKLFSFGGLGIWTLIDVLLIGVGYVG PADGSLYI	1						CTYGKPVTFDCAVKPSVTCVDODFKSOKNFII
RYPANCTVR\(\text{DHVHCLGNRTFPKMLYCNWT}\) GGYKWVYGLWTHHPR\(\text{WQFQVGPVGPVGFXSGLFSFGGLGIWTLIDVLLIGVGYVGPVGPADGSLYI\) 595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGS\(\text{NNSPASASRV\)}\) SPASASRV\(\text{AGNIGARHHT\}\)Q\(\text{UFVLVQMRVHYVGQDGLDL\}\) NLMHPPRSPKVLGLQA 596 1946 A 4513 3 1674 HASDHL\(\text{PMFTLVNETLLK\)QKQFFEEKRFKLD HSVSSTNGHR\(\text{WQIFQDWLGTDQDNLDLANV\}\) NLMLELLV\(\text{UKKK\}\)QLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSSHSSIIDSTEVSQPPGFSGSSQTKK\)QPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KF\(\text{TYTYNSVPLATLSYASDLYNGSQYKSLV\}\) FEFDRDCDYFAIAGVIKKKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHK\(\text{LLASS\}\)DYEGTVIL\(\text{WDGFTGQRSKVYQEHEKRCWSV\}\) DFNLMDPKLLASS DYEGTVIL\(\text{WDGFTGQRSKVYQEHEKRCWSV\}\) DFNLMDPKLLASSDDAKVKL\(\text{WSTNLDNSV\}\) ASIEAKANVCCVKFSPSSRYHL\(\text{AFGCADHCV\}\) HYYDLRNTK\(\text{QPIMVFKGHRKAVSYAKFVSG\}\) EEIVSASTDSQLKL\(\text{WNVGKPYCLRSFKGHIN\}\) EKIVF\(\text{VGLSSNGDYIACGSENNSLYLYYKGLS\}\) KTLLTFKPDTVKSVLDKDRKEDDTNEFVSAV C\(\text{WRALPPGESNVLIAANS\}\)QGTIKVLELV 597 1947 A 4518 536 824 RSLALSFGLECSGMISAHCNL-HLLGSSDPPTS ASQVAEITSVRHTWLIFCNLGQMGFHHVGE\) QAGLELLTSWDPAILPSQSAGIIGMSPHA\(\text{WPP}\}\) 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\(\text{CRPGAYFFSF}\)	1 1						
GGYKWVYGLWLLRHHPRWGLGADRF\YLGP							
S95 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRV/AGNIGARHHTQQIFVLLVQMRVH YVGQDGLDLL/NLMIHPPRSPKVLGLQA 596 1946 A 4513 3 1674 HASDHLYPNFLVNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEASHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLLS KFIRTYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMPFKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANSQGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASVAVEITSVRHTWLIFCILGQMGFHIVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR/CRLPGAYFFSF	1 1						
595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRV/AGNIGARHHTQQIFVLLVQMRVH YVGQDGLDLL/NLMHPPRSPKVLGLQA 596 1946 A 4513 3 1674 HASDHLYPNFLVNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KFNTRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDFKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYVCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLLAANSQGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCNLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFSF							
595 1945 A 4512 533 264 FFFKMESYSVARLECSGAISAPCNLHLLGSNN SPASASRV/AGNIGARHHTQQIFVLLVQMRVH YVGQDGLL/NLMHPPRSPKVLGLQA 596 1946 A 4513 3 1674 HASDHLYPNFLVNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRKRLTAHFE DLEQCYFSTRMSRISDDSKTASQLDEFQECLS KFTRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPVCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANSQGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLJGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFSF	l						
SPASASRV/AGNIGARHHTQQIFVLLVQMRVH YVGQDGLDLL/NLMIHPPRSPKVLGLQA 596 1946 A 4513 1674 HASDHLYPNFLVNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KPITRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTYKSVLDKDRKEDDTNEFVSAV CWRALPPGESNVLIAANSQGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	595	1945	A	4512	533	264	
996 1946 A 4513 3 1674 HASDHLYPNFLVNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFALAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKILASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCCLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF			••		555	201	
1946 A 4513 3 1674 HASDHLYPNFLVNELILKQKQRFEEKRFKLD HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FFFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							ANCOUCH DE L'AIL PARABBBEBRAT OF VY
HSVSSTNGHRWQIFQDWLGTDQDNLDLANV NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLS KFATRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANSQGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	596	1946	A	4512	3	1674	
NLMLELLVQKKKQLEAESHAAQLQILMEFLK VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIDSTEYSQP PGFSGSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS KFYTRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYVCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANSQGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF			T.	4010	,	10/4	
VARRNKREQLEQIQKELSVLEEDIKRVEEMS GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECLLS KFATRYNSVRPLATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNILLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANSQGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF						İ	NI M ELI VOLVEVOI E FERTI LA COLO COLO
GLYSPVSEDSTVPQFEAPSPSHSSIIDSTEYSQP PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQECULS KFATRYNSVRPLVATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHLAFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF					ł	Ì	
PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE DLEQCYFSTRMSRISDDSRTASQLDEFQEC\LS KF\TRYNSVPPLATILSYASDLYNGSQYKSLV FEFDDCDYFAIAGVTKKKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGT\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF					1		
DLEQCYFSTRMSRISDDSRTASQLDEFQECUS KFYTRYNSVRPLVATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF				-	i		
KF\TRYNSVRPL\ATLSYASDLYNGSQYKSLV FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNILLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCILGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF				1			PGFSGSSQTKKQPWYNSTLASRRKRLTAHFE
FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYVCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF							
VDIHYPENEMTCNSKISCISWSSYHKNLLASS DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYVCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANSQGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		. [1	·		
DYEGTVILWDGFTGQRSKVYQEHEKRCWSV DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVQKPYCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		1	ł	ŀ			FEFDRDCDYFAIAGVTKKIKVYEYDTVIQDA
DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYVCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		ļ	-			I	VDIHYPENEMTCNSKISCISWSSYHKNLLASS
DFNLMDPKLLASGSDDAKVKLWSTNLDNSV ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYVCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		ļ					DYEGTVILWDGFTGQRSKVYQEHEKRCWSV
ASIEAKANVCCVKFSPSSRYHI.AFGCADHCV HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKPYCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGT\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCNLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		1		1			
HYYDLRNTKQPIMVFKGHRKAVSYAKFVSG EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGTIKVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCRLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		Ì			ļ		
EEIVSASTDSQLKLWNVGKP\YCLRSFKGHIN EKNFV\GLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGT\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFC\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		i		1	ļ		
EKNFVGLASNGDYIACGSENNSLYLYYKGLS KTLLTFKFDTVKSVLDKDRKEDDTNEFVSAV CWRALPDGESNVLIAANS\QGT\KVLELV 597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFC\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF				l			
S97 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCNLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF				ļ			•
597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCNLGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF			- 1	ŀ	1		
597 1947 A 4518 536 824 RSLALSPGLECSGMISAHCNLHLLGSSDPPTS ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	ĺ		ſ		1	-	
ASQVAEITSVRHHTWLIFCI\LGQMGFHHVGE QAGLELLTSWDPAILPSQSAGIIGMSPHAWPP 598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	597	1947		4518	536	824	
598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF	٠,,	-27'	^	7210	JJ0	024	
598 1948 A 4524 1 384 FDTEFVNIGGDFDAAAGVFR\CRLPGAYFFSF		1		ļ	1		
TOTAL VIOLOTIAN OF THE PROPERTY OF THE PROPERT	600	1040	 _	4526		204	
TLGKLPRKTLSVKLMKNRDEVQAMIYDDGSS	סצנ	1745	A	4524	1	584	
						1	TLGKLPRKTLSVKLMKNRDEVQAMIYDDGSS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ . ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion RRREMQSQSVMLALRRGDAVWLLSHDHDG YGAYSNHGKYITFSGFLVYPDLAPAAPPGLG ASELL
599	1949	A	4526	366	776	MGQPAPYAEGPIQGGDAGELCKCDFLVFTSP NPEAVCEAGTPAMFQTAWRQMESCSI/AQAG VQWRDPGSLHPPPLGFKRFSCLSLPSSWDYK HAPPHPANFCIFSRDQVSPCWPGWSRSLDLVI PPPWLPKVLGLQA
600	1950	A	4529	776	334	FFFETESCYVAQAGVQWCDLCSLQAPPPG\SS DPPASASRVAGTTGARHHTQI.IFVFLVETGFH \MLARDGLKLLTSSDPPASASQSSWDYRREPP RLANFFVFLVETGSRYVAQAGVQWLFTGAIP LLISTGVLTCSVSDLGRFTPP
601	1951	A	4533	1460	403	HEVQESIHFLESEFSRGISDNYTLALITYALSS VGSPKAKPALNMLTWRAEQEGGMQFWVSSE SKLSDSWQPRSLDIEVAAYALLSHFLQFQTSE GIPIMRWLSRQRNSLGGFASTQDTTVALKALS EFAALMNTERTNIQVTVTGPSSPSPVKFLIDT HNRLLLQTAELADGTANGSV/SISANGFGFAI CQLNVVYNVKASGSSRRRRSIQNQEAFDLDV AVKENKDDLNHVDLNVCTSFSGPGRSGMAL MEVNLLSGFMVPSEAISLSETVKKVEYDHGK LNLYLDSVNETQFCVNIPAVRNFKVSNTQDA SVSIVDYYEPRRQAVRSYNSEVKLSSCDLCSD VQRLPSL
602	1952	A	4540	1963	295	MRAPGRPALRPLPLPPLLLLLSSPWGRAVPC VSGGLPKPANITFLSINMKNVLQWTPPEGLQG VKVTYTVQYFIYGQKKWLNKSECRNINRTYC DLSAETSDYEHQYYAKVKAIWGTKCSKWAE SGRFYPFLETQIGPPEVALTTDEKSISVVLTAP EKWKRNPEDLPVSMQQIYSNLKYNVSVLNT KSNRTWSQCVTNHTLVLTWLEPNTLYCVHV ESFVPGPPRRAQPSEKQCARTLKDQSSEFKAK IIFWYVLPISITVFLFSVMGYSIYRYIHVGKEK HP\ANLILIYG\NEFDKRFFVPA\EKIV\INFI\TL NIS\DDSKISHQDMSLLGKSSDVSSLNDPQPSG NLRPPQEEEEVKHLGYASHLMEIFCDSEENT\ EGTSFTQQESLSRTIPPDKTVIEYEYDVRTTDI CAGPEQELSLQEEVSTQGTLLESQAALAVL GPQTLQYSYTPQLQDLDPLAQEHTDSEGPEE EPSTILVDWDPQTGRLCIPSLSSFDQDSEGCE PSEGDGLGEEGLLSRLYEEPAPDRPPGENETY LMQFMEEWGLYVQMEN
603	1953	A	4543	3	600	YSAVEFVEQASGISDWWNPALRKRMLSDSGL GMIAPYYEDSDLKDLSHSRVLQSPVSSEDHAI LQAVIAGDLMKLIESYKNGGSLLIQGPDHCSL LHYAAETGNGEIVKYILDHGPSELLDMADSE TGETALHKAACQRNRAVCQLLVDAGASLRK\ TDSKGKTPQERAQQA\GDPDLAA/YTIESRQN YKVIGHEDLETAV
604	1954	A	4548	3	938	QDNKVQNGSLHQKDTVHDNDFEPYLTGQAN QSNSYPSMSDPYLSSYYPPSIGFPYSLNEAPW STAGDPPIPYLTTYGQLSNGDHHFMHDAVFG QPGGLGNNIYQHRFNFFPENPAFSAWGTSGS QGQQTQSSAYGSSYTYPPSSLGGTVVDGQPG FHSDTLSKAPGMNSLEQGMVGLKIGDVSSSA VKTVGSVVSSVALTGVLSGNGGTNVNMPVS KPTSWAAIASKPAKPQPKMKTKSGPVMGGG LPPPPIKHNMDIGTWDNKGPVPKAPVPQQAP

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ļ		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		l		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	}		l	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			1	peptide		/=possible nucleotide deletion, \=possible
	<u> </u>			sequence	<u> </u>	nucleotide insertion
			1			SPQAAPQPQQVAQPLPAQPPALAQPQYQSPQ OPPO
605	1955	A	4553	2	2304	QPPQ
003	1933	^	4333	 	2304	ILLQEKRNCLLMQLEEATRLTSYLQSQLKSLC
						ASTLTVSSGSSRGSLASSRGSLASSRG FTDIYGLPQYEKPDAEGSQLLRFDLIPFDSLGR
1		•	} .			DAPFSEPPGPSGFHKQRRSLDTPQSLASLSSRS
· '						SLSSLSPPSSPLDTPFLPASRDSPLAQLADSCE
		1				GPGLGALDRLRAHASAMGDEDLPGMAALQP
i i	}					HGVPGDGEGPHERGPPPASAPVGGTVTLRED
		1				SAKRLERRARRISACLSDYSLASDSGVFEPLT
i .						KRNEDAEEPAYGDTASNGDPQIHVGLLRDSG
						SECLLVHVLQLKNPAGLAVKEDCKVHIRVYL
1		1		1		PPLDSGTPNTYCSKALEFQVPLVFNEVFRIPV
						HSSALTLKSLQLYVCSVTPQLQEELLGIAQIN
1						LADYDSLSEMQLRWHSVQVFTS\LNHQGRGR
İ '						LGVQERAPPGTLHTPSPSPA/STDAVTVLLAR
1		1		l		TTAQLQAVERELAEERAKLEYTEEEVLEMER
						KEEQAEAISERSWQADSVDSGCSNCTQTSPPY
						PEPCCMGIDSILGHPFAAQAGPYSPEKFQPSPL
			[KVDKETNTEDLFLEEAASLVKERPSRRARGSP
		1				FVRSGTIVRSQTFSPGARSQYVCRLYRSDSDS
						STLPRKSPFVRNTLERRTLRYKQSCRSSLAEL
		i				MARTSLDLELDLQASRTRQRQLNEELCALRE
1					1	LRQRLEDAQLRGQTDLPPWVLRDERLRGLLR
	,					EAERQTRQTKLDYRHEQAAEKMLKKASKEI
]				, .		YQLRGQSHKEPIQVQTFREKIAFFTRPRINIPPL
606	1956	A	4555	3429	776	PADDV PGSGPGPAPFLAPVAAPVGGISFHLQIGLSREP
000	1530	Λ.	(0,00	J447	, ,0	VLLLQDSSGDYSLAHVREMACSIVDOKFPEC
						GFYGMYDKILLFRHDPTSENILQLVKAASDIQ
						EGDLIEVVLSASATFEDFQIRPHALFVHSYRA
1						PAFCDHCGEMLWGLV\RQGLKCEGCGLNYH
						KRCAFKIPNNCSGVRRRRLSNVSLTGVSTIRT
						SSAELSTSAPDEPLLQKSPSESFIGREKRSNSQ
						SYIGRPIHLDKILMSKVKVPHTFVIHSYTRPTV
					•	CQYCKKLLKGLFRQGLQCKDCRFNCHKRCA
					'	PKVPNNCLGEVTINGDLLSPGAESDVVMEEG
						SDDNDSERNSGLMDDMEEAMVQDAEMAMA
						ECQNDSGEMQDPDPDHEDANRTISPSTSNNIP
						LMRVVQSVKHTKRKSSTVMKEGWMVHYTS
						KDTLRKRHYWRLDSKCITLFQNDTGSRYYKE
]						IPLSEILSLEPVKTSALIPNGANPHCFEITTANV
						VYYVGENVVNPSSPSPNNSVLTSGVGADVAR
						MWEIAIQHALMPVIPKGSSVGTGTNLHRDISV
	i					SISVSNCQIQENVDISTVYQIFPDEVLGSGQFGI
						VYGGKHRKTGRDVAIKIIDKLRFPTKQESQLR
						NEVAILQNLHHPGVVNLECMFETPERVFVVM
						EKLHGDMLEMILSSEKGRLPEHITKFLITQILV
						ALRHLHFKNIVHCDLKPENVLLASADPFPQV
					!	KLCDFGFARIIGEKSFRRSVVGTPAYLAPEVL
1						RNKGYNRSLDMWSVGVIIYVSLSGTFPFNED
						EDIHDQIQNAAFMYPPNPWKEISHEAIDLINN
						LLQVKMRKRYSVDKTLSHPWLQDYQTWLDL
						RELECKIGERYITHESDDLRWEKYAGEQGLQ
L					4400	YPTHLINPSASHSDTPETEETEMKALGERVSIL
607	1057	_	1562	1		
607	1957	A	4563	1	4499	SRPWWLRASERPSAPSAMAKRSRGPGRRCLL
607	1957	A	4563	1	4499	ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT
607	1957	A	4563	1	4499	ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT TVRCMHLLLEAVPAVAPQTSILDLRFNRIREI
607	1957	A	4563	1	4499	ALVLFCAWGTLAVVAQKPGAGCPSRCLCFRT

OPO TO	050	1.64	000			
SEQ II		Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nuci-	peptide	ļ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Laucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	- 1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1			peptide		/=possible nucleotide deletion, \=possible
		<u> </u>		sequence		nucleotide insertion
	ł			T -		HFNQIETLDPDSFQHLPKLERLFLHNNRITHL
		1	1	1		VPGTFNHLESMKRLRLDSNTLHCDCEILWLA
	j	1	1]		DLLKTYAESGNAQAAAICEYPRRIQGRSVATI
			i i			TPEELNCERPRITSEPQDADVTSGNTVYFTCR
	1.	Į.				AEGNPKPEUWLRNNNELSMKTDSRLNLLDD
	l l		1			GTLMIQNTQETDQGIYQCMAKNVAGEVKTQ
	ı	ĺ	1	ĺ	[EVTLRYFGSPARPTFVIQPQNTEVLVGESVTL
]		1	ŀ	1	ECSATGHPPPRISWTRGDRTPLPVDPRVNITPS
	i	[1			GGLYIQNVVQGDSGEYACSATNNIDSVHATA
	1	ł	1			FIIVQALPQFTVTPQDRVVIEGQTVDFQCEAK
	ı	ļ	1		ļ	GNPPPVIAWTKGGSQLSVDRRHLVLSSGTLRI
		1			ł	SGVALHDQGQYECQAVNIIGSQKVVAHLTVQ
		1			[PRVTPVFASIPSDTTVEVGANVQLPCSSQGEP
	1	'			[EPAITWNKDGVQVTESGKFHISPEGFLTINDV
		l	1	Ī	1	GPADAGRYECVARNTIGSASVSMVLSVNVPD
						VSRNGDPFVATSIVEALATVDRAINSTRTHLF
	1	Ì				DSRPRSPNDLLALFRYPRDPYTVEQARAGEIF
	1	l	1			ERTLQLIQEHVQHGLMVDLNGTSYHYNDLVS
						PQYLNLIANLSGCTAHRRVNNCSDMCFHQKY
		İ	1			RTHDGTCNNLQHPMWGASLTAFERLLKSVY
]			
			1			ENGFNTPRGINPHRLYNGHALPMPRLVSTTLI
	1		1			GTETVTPDEQFTHMLMQWGQFLDHDLDSTV
	1		1 1			VALSQARFSDGQHCSNVCSNDPPCFSVMIPPN
	1.	ł				DSRARSGARCMFFVRSSPVCGSGMTSLLMNS
	I.	ļ				VYPREQINQLTSYIDASNVYGSTEHEARSIRD
	1	ĺ				LASHRGLLRQGIVQRSGKPLLPFATGPPTECM
	1					RDENESPIPCFLAGDHRANEQLGLTSMHTLW
	j	J	J l	•		FREHNRIATELLKLNPHWDGDTTYYETRKIVG
		,	1	٠		AEIQHITYQHWLPKILGEVGMRTLGEYHGYD
	1		1 1	ļ		PGINAGIFNAFAT\AAFRFGHTLVNPLLLPGLD
	į	1	1			ENFQPIAQDHLPLHKAFFSPFRIVNEGGIDPLL
	1	İ	1 1			RGLFGVAGKMRVPSQLLNTELTERLFSMAHT
	1	j	1 }			VALDLAAINIQRGRDHGIPPYHDYRVYCNLS
			1			AAHTFEDLKNEIKNPEIREKLKRLYGSTLNID
	i	1	1 1			LFPALVVEDLVPGSRLGPTLMCLLSTQFKRLR
	j	l	!!!			DGDRLWYENPGVFSPAQLTQIKQTSLARILCD
		[ĺ	İ	NADNITRVQSDVFRVAEFPHGYGSCDEIPRVD
						LRVWQDCCEDCRTRGQFNAFSYHFRGRRSLE
			<u> </u>	ļ		FSYQEDKPTKKTRPRKIPSVGRQGEHLSNSTS
			1 1		!	A\FSTRSDASG\TNDFQRVCSWEMQKTITDLR
		1	1		l	TQIKKLESR\LSTTECVDAGGESHANNTKWK
	1			ſ		KDACTICECKDGQVTCFVEACPPATCAVPVNI
		L	L i			PGACCPVCLQKRAEEKP
608	1958	A	4566	354	1135	FSFLC/GVSGRLGLDSEEDYYTPQKVDVPKAL
	1		- - -		4	IIVAVQCGCDGTFLLTQSGKVLACGLNEFNKL
				. 1		GLNQCMSGIINHEAYHEVPYTTSFTLAKQLSF
			1	1		YKIRTIAPGKTHTAAIDERGRLLTFGCNKCGQ
			} }	ì	ļ	LGVGNYKKRLGINILLGGPLGGKQVIRVSCGD
				1		EFTIAATDDNHIFAWGNGGNGRLAMIPTERP
				ľ		HGSDICTSWPRPIFGSLHHVPDLSCRGWHTILI
				ļ	ļ	VEKVLNSKTIRSNSSGLSIGTVFQSSSPGGGGE
				[GGPDAW
609	1959	A ·	4567	1	412	
-02	""	Α.	1307	•	712	FFFFETESRSVAQAGVQWRDLGSLQAPPPGFT
	1			l		PFSCLSLPSSWDYRRPPLRPANFFVFLVETGF
] !		ļ j	}		HRFSRDGLDLLT/S/GDPPASASQSAGITGVSH
				ŀ	ļ	RARPRINLRNVIYSFAVTYCLNYISLAMSSTL
610	1000		4570	(02	465	KLSFHVLSGS
610	1960	A	4570	697	467	ECRGVISAH\CCTLCLPSSSDSASAF\RVARTT
				1	İ	GTCDYAQLIFAFLVEMGFHHVGQDGLHLL/N LVIRPPRPPKVLGLQA

			Lono	- T	1 70 - 41 - 4 - 4 4	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of peptide	поц	in NO.	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl- eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	nence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	Lucico		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
псисс		1	^**	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	· •	/=possible nucleotide deletion, \=possible
		1	1	sequence	1	nucleotide insertion
611	1961	A	4571	25	1396	ADPHTTVIRFFPAASATKRVLPPVLRVSSPRT
		ļ				WNPNVPESPRIPAPRLPKRMSGAPTAGAALM
	İ]		ļ)	LCAATAVLLSAQGGPVQSKSPRFASWDEMN
				1		VLAHGLLQLGQG\CANT\GAHPQSAERAGA\R
	Į.	1				LSACGSACQGTEGSTDLPLAPESRVDPEVLHS
	1			1	Ì	LQTQLKAQNSRIQQLFHKVAQQQRHLEKQHL
	[ĺ	1	1		RIQHLQSQFGLLDHKHLDHEVAKPARRKRLP
	1					EMAQPVDPAHNVSRLHRLPRDCQELFQVGER
	Ì	ì		1		QSGLFEIQPQGSPPFLVNCKMTSDGGWTVIQR
Ì	1	1	1			RHDGSVDFNRPWEAYKAGFGDPHGEFWLGL
ļ)		1	ļ		EKVHSITGDRNSRLAVQLRDWDGNAELLQFS
	ł		}		1	VHLGGEDTAYSLQLTAPVAGQLGATTVPPSG LSVPFSTWDQDHDLRRDKNCAKSLSGGWWF
]	ĺ		1			GTCSHSNLNGQYFRSIPQQRQKLKKGIFWKT
ŀ	1	1	ļ	Ì	1	WRGRYYPLQATTMLIQPMAAEAAS
	1000		4575	162	3	FFFETESRSVAQAGVQWRDLSSLQPPPPG\SR
612	1962	Α	43/3	102	'	GSPASASPVAGITGTRHHRTRG
613	1963	A	4584	687	321	PLAQRRPFLWVTVKTNGHIWGSSTYPHFWGS
013	1903	^	4304	007	321	SNS/PASASQVAGIPNARHQARIIFVFLVEPRF
		}		1		HHVGRAGLGFL/NLAICLPQHPKVLGLQACN
		1		1	•	LNIKPHPAHKYISMIQFNVHFMCMSVHIYI
614	1964	A	4589	727	299	PGSAQSAQRGRGRRRARAGSATQITMYSFMG
] 014	1501	1	1307	1		GGLFCAWVGTILLVVAMATDHWMQYRLSGS
		1		1	1	FAHOGLWRYCLGNKCYLQTDSIAYWNATRA
		1				FMILSALCAISGIIMGIMAF/GWVAVLMTFFA
]	1	1	1	GIFYMCAYRVHECRRLSTPR
615	1965	A	4590	2	414	TILPEKIQAWAQKQCPQSGEEAVALVVHLEK
1		i	1	1	İ	ETGRLRQQVSSPVHREKHSPLGAAWEVADFQ
			1	,		PEQVETQPRAVSREEPGSLHSGHQEQLNRKR
Ì		1		1	1	ERRPLPKNARPSPWVPALADEWNTLHQEVTT
<u> </u>	J	1				TRLPAGSQEPVKD
616	1966	A	4592	773	488	DFALVAQAGVQWHNLGSPQPLPPGFKRFSCL
		Į				SLPSSWEYRCVPP/RLANFVFLVEMGFLHVGQ
L	1	١	4505	104	478	AGLELPTSGDPPALASQSAGITGVTTVPSGPG XRHGLREPLLERRCAAASSFQHSSSLGRELPY
617	1967	В	4595	84	4/8	DPVDTEGFGEGGDMQERFLFPEYILDPEPQPT
1	1	1	1	Ì		REKQLQELQQQEEEERQRQQRREERRQQNL
1	1			ł		RARSREHPVVGHPDPALPPSGVNCSGCGAEL
		1	1	1		HCQDAR*
618	1069	+	4596	2945	1188	ARSRNSARGVYGMCVDTLFLCFLEDLERNDG
618	1968	A	1330	1	1100	SAERPYFMCSTLKKPLARRCFPAIHAYKGVL
1	[1	1	1		MVGNETTYEDGHGSRKNITDLVEGAKKANG
1	1	1	1	ì		VLEARQLAMRIFEDYTVSWYWIIIGLVIAMA
1.		-	1		<mark> </mark> ,	MSLLSIILLHLLAGIMGWVMIIMEI\SELGYRIF
1			1	ı		HCYMEYSRLRGEAGSDVSLVDLGFQTDFRV
			1	1		YLHLRQTWLAFMIILSILEVIIILLLIFLRKRILI
				1		AIALIKEASRAVGYVMCSLLYPLVTFFLLCLCI
1		1		1	1	AYWASTAVFLSTSNEAVYKIFDDSPCPFTAKT
				1		CNPETFPSSNESRQCPNARCQFAFYGGESGYH
		1		1	1	RALLGLQIFNAFMFFWLANFVLALGQVTLAG
1	1	1		1	Í	AFASYYWALRKPDDLPAFPLFSAFGRALRYH
1	1			1	1	TGSLAFGALILAIVQIIRVILEYLDQRLKAAEN
1		1	1			KFAKCLMTCLKCCFWCLEKFIKFLNRNAYIM
1	1		1	1	1	IAIYGTNFCTSARNAFFLLMRNIIRVAVLDKV
	1	}	}	j	1	TDFLFLLGKLLIVGSVGILAFFFFTHRIRIVQDT
1			ł	1	1	APPLNYYWVPILTVIVGSYLIAHGFFSVYGMC VDTLFLCFLEDLERNDGSAERPYFMSSTLKKL
1		ļ		1		LNKTNKKAAES
619	1000	+-	4601	12	357	RTSVEPYILGEF/RKLSNNTKVVKTEYKATEY
i DIY	1969) A	4001	1 -		***************************************

SSQ II		LOPOID	1 N d - 4	LOPA	I D 3: - *	1	74 41
muchoide muchoide	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Sequence			поа				
Seq			ì				
mence				i			
amino acid residue of peptide residue of peptide sequence T=Threonine, V=Valine, V=Typtophen, Y=Typtophen		Baice		1		1	
Pepsible peptide Peptide Pepsible Pe	uence		ł	714			
Peptide	[
Sequence	1					Sequence	
1976 A 4629 249 3				İ			
	626	1976	A	4629		3	
ASASQVAGIAGTHE A A635 1 301				1			
1977 A	(İ	ĺ	1		ĺ	1
	627	1977	A	4635	1	301	
ARAYL	İ						
ARAYL				-			PGGAG*PRPPDLRGPPGLAPPQGGNNGGDPP
TYGNNAFILYTLYTRIVILLF*HV*PA YFQPSK NKTAKINGN*RFFI.FI.VCVILL*AFI.HIGIFIAN							
NKTAKINCN-RPFILI-VCYLL*AFILHIOIFIANF	628	1978	Α	4648	1357	782	KLFSSQRLFGPHIQAINPSFLLLSFFPS*LLAMR
	1			1 .			TVGNNAFILVFLVYRIVLLLF*HV*PAYFQPSK
	l		1				NKTAKINCN*RPFI.FI.VCYLL*AEI.HIGIFIANF
SYKSFAVIIFFUNITRFFSGF						ļ	
1979 A 4660 18 999		ł	Ì				
KVKTVCQPLMKEMLKRFQVAVNLAEDTAH PKL VFSQEGRYYKNTASASSWVPSSAWNY AGWRNPQKTAFVERFQHLSCVLGKNVFTSG KHYWEVESRDSLEVAVGVCREDVMGITDRS KHYWEVESRDSLEVAVGVCREDVMGITDRS KHYWEVESRDSLEVAVGVCREDVMGITDRS KHYWEVESRDSLEVAVGVCREDVMGITDRS KHYWEVESRDSLEVAVGVCREDVMGITDRS KHYWEVESRDSLEVAVGVCREDVMGITDRS KHYWEVESRDSLEVAVGVCREDVMGITDRS KHYWEVESRDSLEVAVGVCREDVMGITDRS KHYWEVESRDSLEVAVGVCREDVMGITDRS KHYWEVESRDSLEVAVGVCREDVMGITDRS KHYWEVESRDSLEVAVGVCREDVMGITDRS KHYWEVESPLASLVIPPVLGFTGTYDRX*G FSSPDQNSFPRVUSITDHPWLFCFSCLYPG WSIFWVSLTVPFGICPLCASQEAVPWEVGLA NGDGTGMFPRFREWEIFL MSGGTGMFPRFREWEIFL MSGGTGMFPRFREWEIFL MSGGTGMFPRFREWEIFL MFFFFETESHSVAQAGMQWRNLGSLPAPPPGF TFFFCLSLLNGWDVRRPPPHLAMFPLLVETG FHEDVGQDGLDLLTS*STPSASQSAEITGVSHC TRLKKIRFAKGHVEFFFESHVE TFFFGTDDEHEECTVQEYSAGKNTCLRPGAV AHTCNPCTLGGRGRWITYGSAQCHQUPQ NPVFLERRPRALHSSPGLTTQRIL WAQQLWV GAGSTGCSRGFREGVFREG AHTCNPCTLGGRGRWITYGSAQCHQUPQ NPVFLERRPRALHSSPGLTTQRIL WAQQLWV GAGSTGCSRGFREGVFREG GAGSTGCSRGFREGGVFREG GASTGCSRGFREGGVFREG GASTGCSRGFREGGVFREG GASTGCSRGSRGFREGGVFREG GASTGCSRGSALLQAPVLTWTYSELSRSVATATTC GASTGCSRGSALLQAPVLTWTYSELSRSVATATTC GASTGCSRGSALLQAPVLTWTYSELSRSVATATTC GASTGCSRGSALLQAPVLTWTYSELSRSVATATTC GASTGCSRGSALLQAPVLTWTYSELSRSVATATTC GASTGCSRGSALLQAPVLTWTYSELSRSVATATTC GASTGCSRGSALLQAPVLTWTTSTSGRG GAPFGCSSGSSALLQAPVLTWTTSTSGRG GASTGCSRGSALLQAPVLTWTTSTSGRG GASTGCSRGSALLQAPVLTWTTSTSGRG GASTGCSRSSALLQAPVLTWTTSTSGRG GASTGCSRSSALLQAPVLTWTTSTSGRG GASTGCSRSSALLQAPVLTWTTSTSGRG GETDSSEFEVRRYQDFLWLKSGRGSCSALLAPVLTWTTSTSGRG GASTGCSRSSALLQAPVLTWTTSTSGRG GASTGCSRSSALLQAPVLTWTTSTSGRG GASTGCSRSSALLQAPVLTWTTSTSGRG GASTGCSRSSALLQAPVLTWTTSTSGRG GASTGCSRSSALLQAPVLTWTTSTSGRG GASTGCSRSSALLQAPVTTTSTTSTTSTTSTTSTTTTTTTTTTTTTTTTTTTTT			<u> </u>				
PKLVFSQEGRYYKITASASSWPVSAMYTS AGWRNPOKTAPVERFQHLSCVLGKNVFTSG KHYWEVESRDSLEVAVGVCREDVMGITDRS KMSPDVGIWAIYWSAAGYWPLIGFGTPTQQ EPALHENGYVLDRGTGNNSFYSAVDGVHLH TFSCSSVSRLRPFFWLSPLASLVIPPYTDRK*G FSSPDQNSFPVQLRDTHPWALFCPSCLYPG WSIFWVSLTVPFGICPLCASQEAVPWEVGLA NGDGTGNFPRRFWEIFL TFFFETEISSWAQAGMQWRNLGSLFAPPPGF TFFFETEISSWAQAGMQWRNLGSLFAPPPGF TFFFETEISSWAQAGMQWRNLGSLFAPPPGF TFFFETEISSWAQAGMQWRNLGSLFAPPPGF TFFFETEISSWAQAGMQWRNLGSLFAPPPGF TFFFETEISSWAQAGMQWRNLGSLFAPPPGF TFFFETEISSWAQAGMQWRNLGSLFAPPPGF TFFFETEISSWAQAGMQWRNLGSLFAPPPGF TFFFTETEISSWAQAGMQWRNLGSLFAPPPGF TFFFTETEISSWAQAGMQWRNLGSLFAPPPGF TFFFTETEISSWAQAGMQWRNLGSLFAPPPGF TFFTTDSASQSAEITGVSLF TFFTTETSASQSAEITGVSLF TFFTTETSASQSAEITGVSLC A 4674 953 614 TPRGTDDEHEECTVQEYSAGKNTCLRPGAV AHTCNPCTLGGGRWITY-SGSVQDQDQPFTWQ NPVFLERRPALHSSPGLTTQRLWAQGLWV GAGSTGCSRGFGGVFFEGUWAQGLWV GAGSTGCSRGFRGGVFFEGUWAQGLWV GAGSTGCSRGFRGGGVFFEGU SAGLAPVTHWTYSELSRGKVYCTFP A 4678 34 314 RSTHASGMISPSFGFMGHLLRLEFEILPSTPNP *LPSYQGEAAGSSLISHLQTTSPDLKGVYCTFP A 50LAPVTHWTYSELSRSPVATATTC GASTAGVSRGPGGANGESPGG GAPFPGSSGSSALLQAEVLDLDEDEDDLEVFS KDASLMDMNSFSPMATTSLSMINQIKFEDEP DLKDLFITVDEPESHVTITETTTYRITKTSRG EFDSSEFEVERRYQDFLWIKGKLEEAPPTLII PPLEFKIVKGMVERRNDDFIETRRKALHKEL NRIADHPTLTFNEDFKIRTLAQAWELSHKKQ GFGLLSRMQTVRAVASSMGVKNRPEEFM EMNNFIELFSQKINLDKISQRIYKEREFYPDE MKEYGPHIL WSASEEDLYDTLKDVASCURC CKATEKRMSGLSSALLPVVHEVVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVEVLANALKADWERWKQNM QNDIKLAFTDMAEENHIYYEQCLATWESFLT SOTNILHEEASBOKP GAGSTGCSRGAA*GRKILQFCVH WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGGSCAA*GRKRILQFCVH WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCAAA*GRKRILQFCVH WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGGSCAAA*GRKRILQFCVH WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCAAA*GRKRILQFCVH WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCAAA*GRKRILQFCVH WISKAVYKHREMCGLTSTGRKSHGLEKDRM F	629	1979	A	4660	18	999	HHELHTLELLQNPKEVLTRSEIQDVNYSLEAV
AGWRNPQKTAFVERPQHLSCVLQKNVFTSG KHYWEVESRDSLEVAVGVCREDVMGITDRS KMSPDVGIWAIYWSAGVWPLIGPFGTPTQQ EPALHRVGVYLDRGTGNYSFYSAVDGVHLH TFSCSSVSSLRFFFWLSPLASLASLVPPVTDRTR*G FSSPDQNSFPVVQLRDTHPWALFCPSCLYPG WSIFWVSLTVPFGICPLCASQEAVPWEVGLA MQDGTGNPPRRFWEIL	1		}				
KHYWEVÉSRÖLEVAVĞVCREDYMĞITDRS KMSPDVĞIWAİYWSAAĞVÜRLĞIPĞTİTQQ EPALHRYĞVYYDRĞIĞNVSFYSAVDĞVHLH TFSCSSVSRLRFFWLSPLASLVIPPVTÜRKĞĞ FSSPDQNSPFVVQLBYBWAJFÇPSÇLYPĞ FSSPDQNSPFVVQLBYBWAJFÇPSÇLYPĞ WSIFWVSLTVPFĞICPLCASQEAVPWEVQLA NGDĞTĞNFPRFWEIFL			l				
KMSPDVGIWAIYWSAAGVWPLIGFGTPTQQ			ĺ	İ			
	İ		ľ	1		i	
TFSCSSVSRLRPFFWLSPLASLVIPPVTDRX*G FSSPDQNSFPVVQLRDTHPWALFCPSCLYPG WSIFWVSLTVPFGICPLCASQEAVPWEVGLA NGDGTGNFPRFWEIFL							1 = ** 1
FSSPDQNSFPVQLRDTHPWALFCPSCLYPG WSIFWVSLTVPFGICPLCASQEAVPWEVGLA NGDGTGNFPRRFWEIFL							
WSIFWYSLTYPFGICPLCASQEAVPWEVGLA NGDGTGNFPRRFWEIFL							
	[
1980 A 4669 2 358							
TPFFCLSLLNGWDYRRPPPHLANFFVLLVETG	630	1980	A	4669	2	358	
FFIDVGQDGLDLLTS*STPSASQSAEITGVSHC TRLKKIRFAKGHVEFFFESHVE		1700		'00'	_	""	
TRLKKIRFAKGHVEFFESHVE					•		
1981	f			1			
AHTCNPCTLGGRGRWÌT*GSGVQDQPGPTWQ NPVFLERRPRALHSSPGLTTQRILWAQGLWV GAGSTGCSRGPRGEGVFREG 632 1982 A 4678 34 314 RSTHASGMSPSFGFMGHLLRLEFEILPSTPNP *LPSYQGEAAGSSLISHLQTFSPDLKGVYCTFP ASGLAPVPTHWTVSELSRSPVATATFC 633 1983 A 4696 1 1365 RTLGMEGERRASQAPSSGLPAGGANGESPGG GAPPFGSSGSSALLQAEVLDLDEDEDDLEVFS KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP DLKDLFITVDEPESHVTTIETFTTYRITKTSRG EFDSSEFEVRRRYQDFLWLKGKLEEAHPTLII PPLPEKFIVKGMVERPNDDFIETRRKALHKFL NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHLWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRQQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFIKAIRRNPDTQ WISKAVYKHREMGGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS	631	1981	A	4674	953	614	
GAGSTGCSRGPRGEGVFREG							l •
1982 A 4678 34 314 RSTHASGMISPSFGFMGHLLRLEFEILPSTPNP *LPSYQGEAAGSSLISHLQTFSPDLKGVYCTFP ASGLAPVTHWTVSELSRSPVATATFC ASGLAPVTHWTVSELSRSPVATATFC ASGLAPVTHWTVSELSRSPVATATFC GAPFGSSGSALLQAEVLDLDEDEDDLEVFS KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP DLKDLFITVDEPESHVTTIETFITYRIITKTSRG EFDSSEFEVRRYQDFLWLKGKLEEAHPTLII PPLPEKFIVKGMVERFNDDFIETRRKALHKFL NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS							
*LPSYQGEAAGSSLISHLQTFSPDLKGVYCTFP ASGLAPVPTHWTVSELSRSPVATATFC 633 1983 A 4696 1 1365 RTLGMEGERRASQAPSSGLPAGGANGESPGG GAPFPGSSGSSALLQAEVLDLDEDEDDLEVFS KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP DLKDLFITVDEPESHVTTIETFTTYRITKTSRG EFDSSEFEVRRRYQDFLWLKGKLEEAHPTLII PPLPEKFIVKGMVERFNDDFIETRRKALHKFL NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHLWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS							GAGSTGCSRGPRGEGVFREG
ASGLAPVPTHWTVSELSRSPVATATFC 633 1983 A 4696 1 1365 RTLGMEGERRASQAPSSGLPAGGANGESPGG GAFFPGSSGSSALLQAEVLDLDEDEDDLEVFS KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP DLKDLFITVDEPESHVTTIETFTTYRIITKTSRG EFDSSEFEVRRYQDFLWLKGKLEEAHPTLII PPLPEKFIVKGMVERFNDDFIETRKALHKFL NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEFFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDK VECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS	632	1982	A	4678	34	314	RSTHASGMISPSFGFMGHLLRLEFEILPSTPNP
633 1983 A 4696 1 1365 RTLGMEGERRASQAPSSGLPAGGANGESPGG GAPFPGSSGSSALLQAEVLDLDEDEDDLEVFS KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP DLKDLFITVDEPESHVTTIETTTYRIITKTSRG EFDSSEFEVRRYQDFLWLKGKLEEAHPTLII PPLPEKFIVKGMVERFNDDFIETRRALHKFL NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHTYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFMMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS				}			
GAPFPGSSGSSALLQAEVLDLDEDEDDLEVFS KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP DLKDLFITVDEPESHVTTIETFTYRIITKTSRG EFDSSEFEVRRYQDFLWLKGKLEEAHPTLII PPLPEKFIVKGMVERFNDDFIETRRKALHKFL NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNPPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDFFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLOFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTYHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS				<u></u>			
KDASLMDMNSFSPMMPTSPLSMINQIKFEDEP DLKDLFITVDEPESHVTTIETFITYRITKTSRG EFDSSEFEVRRRYQDFLWLKGKLEEAHPTLII PPLPEKFIVKGMVERFNDDFIETRRKALHKFL NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS	633	1983	Α	4696	1	1365	
DLKDLFITVDEPESHVTTIETFTTYRIITKTSRG EFDSSEFEVRRYQDFLWLKGKLEEAHPTLII PPLPEKFIVKGMVERPNDDFIETRRKALHKFL NRIADHPTLTTNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS							GAPFPGSSGSSALLQAEVLDLDEDEDDLEVFS
EFDSSEFEVRRRYQDFLWLKGKLEEAHPTLII PPLPEKFIVKGMVERFNDDFIETRRKALHKFL NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS				([
PPLPEKFIVKGMVERFNDDFIETRRKALHKFL NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS							
NRIADHPTLTFNEDFKIFLTAQAWELSSHKKQ GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS						•	
GPGLLSRMGQTVRAVASSMRGVKNRPEEFM EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS					į		
EMNNFIELFSQKINLIDKISQRIYKEEREYFDE MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS							
MKEYGPIHILWSASEEDLVDTLKDVASCIDRC CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEFASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS				[
CKATEKRMSGLSEALLPVVHEYVLYSEMLM GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH FPHAIGGSCRAA*RRKTLQFPCYH WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS							
GVMKRRDQIQAELDSKVEVLTYKKADTDLL PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH FPHAIGGSCRAA*RRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS							
PEEIGKLEDKVECANNALKADWERWKQNM QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS				ļ ļ		.]	
QNDIKLAFTDMAEENIHYYEQCLATWESFLT SQTNLHLEEASEDKP 634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS							
SQTNLHLEEASEDKP SQTNLHLEEASEDKP SQTNLHLEEASEDKP SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH SED SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH FRAIGHTAN FPHAIGGSCRAA*RRKTLQFPCYH FRAIGHTAN FPHAIGGSCRAA*RRKTLQFPCYH FRAIGHTAN FPHAIGGSCRAA*RRKTLQFPCYH FRAIGHTAN FPHAIGGSCRAA*RRKTLQFPCYH FRAIGHTAN FPHAIGGSCRAA*RRKTLQFPCYH FRAIGHTAN FPHAIGGSCRAA*RRKTLQFPCYH FRAIGHTAN FPHAIGGSCRAA*RRKTLQFPCYH FRAIGHTAN FPHAIGGSCRAA*RRKTLQFPCYH FRAIGHTAN FPHAIGGSCRAA*RRKTLQFPCYH FRAIGHTAN FPHAIGGSCRAA*RRKTLQFPCYH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RRKTLQFPCH FPHAIGGSCRAA*RKTLQFPCH FPHAIGGSCRAA*RKTLQFPCH FPHAIGGSCRAA*RKTLQFPCH	.						
634 1984 A 4708 421 158 SYWVGEDYTYKFFEVILIDPFHKAIRRNPDTQ WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS					l	l	
WISKAVYKHREMCGLTSTGRKSHGLEKDRM FPHAIGGSCRAA*RRKTLQFPCYH 635 1985 A 4709 42 341 YIKQPDAKERRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS	634	1984	Α	4708	421	158	
635 1985 A 4709 42 341 YIKQPDAKERRRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS							
635 1985 A 4709 42 341 YIKQPDAKERRRTVHWKKETESEASEITIPPST PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS					1	ĺ	
PGVPQAPGHWEDYGRGDNFYLPH*DPGGIVL WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS	635	1985	A	4709	42	341	
WNIFNRMPIARKNITDGEHHEYLIEVPRLFHT SED 636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN°GDLVHCLGGIS					l		•
636 1986 A 4721 2 351 EKPDHFFPEGTSFIHEPRRPN*GDLVHCLGGIS					l		
The last the				L			1
RSTTVTVA+LMQKLNLSMNDAYYIVIMKMSS	636	1986	Α	4721	2	351	
							RSTTVTVA+LMQKLNLSMNDAYYIVIMKMSS

	Q ID); of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nuc	. ,	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cot	ide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq	r	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uen	nce	l	ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1				Ì	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
					residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1					peptide		/=possible nucleotide deletion, \=possible
<u></u>					sequence	ļ	nucleotide insertion
.							ISPNFNSMDQPLDFQRTLGLRSPCYNRVPAQK MYFTTPSNHNAYQVDSVQST
637	, 	1987	A	4726	664	253	NTGLTCSIQRKCGETQLYRREENRLILLLQDH
							LKSESFQVLTLSPRLEFSGLISAHCNLRLPGSS
	i						DSSASSSRAAGITGVHHHAWLIFFFLVETGFL
							HAG*AGLELLTSGDPPASASRSAGITGVSHHA
							RPRETRFL
638	3	1988	Α	4734	24	592	GGMDSRVSGTTSNGETKPVYPVMEKKEEDG
1							TLERGHWNNKMEFVLSVAGEIIGLGNVWRFP
	-						YLCYKNGGGAFFIPYLVFLFTCGIPVFLLETAL
							GQYTSQGGVTAWRKICPIFEGIGYASQMIVIL
	,					1	LNVYYIIVLAWALFYLFSSFTIDLPWGGCYHE
639		1989		4743	1040	699	WNTEHCMEFQKTNGSLNGTSENATSPVIEFW
637	"	1707	A	4/43	1040	עלס	QGLTLLPRMECSATITAHCSLELPGSIDLPTSA S*VARTTGTHHHPWLILVLLL*TWGSYYVAQ
	- 1						AGLELLGSSNLPAAMVSQSAQIIGHDHCAWA
	.				!		TSNHVLYTOEGLRRGKEG
640	, 	1990	A	4771	527	2	GRIDCPHPATVLAQPIFIDACSVLGAYQGAON
				''''	52,	-	WIRRRPCLPSGCLKMNREIGPLOHSLCCPGWS
1							QTPGLKAILLROPPK*LGLOMESHSCPPAWSA
l	}					}	MARSRLTATSASQVQAILLPQPPGTTDSCSPS
							PDHEQQPLSWVLPPPQKDMNPREQQVALGP
L							QAAALPWAVWRNDCFPR
641		1991	Α	4780	16	473	RPSSQCGGIPTGWKKGLAPELSSELSSPPLPAR
Ì	ļ						LQLAASPYFSPSWAECPQPVPAGTHATWCLA
	[RVWARMTPPGPAGIPSHPLPPPPPERSVPIPSP
1							FPARDSGSRQGHSTDRYKHTDAPRDAHRRVP
642	, 	1992	Α	4798	1	487	QRDTDTGVHTGSGTHTHAHTPPEK GYSFRCDIVDYSRSPTALRMARTCWLYYFSK
1 072	١	1//2	Λ.	1770	•	407	FIELLDTIFFVLRKKNSQVTFLHVFHHTIMPW
							TWWFGVKFAAGGLGTFHALLNTAVHVVMY
							SYYGLSALGPAYQKYLWWKKYLTSLQLVQF
l	j						VIVAIHISQFFFMEDCKYQFPVFACIIMSYSFM
L							FLLLFLH
643	3]	1993	A	4799	2	391	LMAFIEMHISGSLVYLKIKTKIYSYFSMLNFLL
							QEIPLSEILRISSPRDFTNISQGSNPHCFEIITDT
							MVYFVGENNGDSSHNPVLAATGVGLDVAQS
1	ĺ						WEKAIRQALMPVTPQASVCTSPGQGKDHSK Q*ASVCTSPGQGKDHSKQ
644		1994	A	4800	488	101	AYPLFAVHPVHTECVAGVVGRAYLLCALFFL
""	'	4774	^	7000	700	101	LSFLGYCKAFRESNKEGAHSSTFWVLLSIFLG
1		-					AVAMLCKEQGITVLVRAATWLGPAFSVCPFP
ľ							SYKDIWGWPCLCGVLHAYIPLLV
645	5	1995	A	4805	458	126	LLWTTVLCQTPARPQSTMIHLGHILFLLLLPV
1							AAAQTTPGERSSLPAFYPGTSGSCSGCGSLSL
[PLLAGLVAADAVASLLIVGAVFLCARPRRSP
		•					AQEDGKVYINMPGRG
646	5 T	1996	A	4817	47	1033	LQGDTWHLSFLSHFSRLHGGVPGRGLLEGNL
l	l						LQPQAPGHDMTSIPFPGDRLLQVDGVILCGLT
	ŀ			ĺ	,	1	HKQAVQCLKGPGQVARLVLERRVPRSTQQC
1							PSANDSMGDERTAVSLVTALPGRPSSCVSVT
l							DGPKF*SSN*KRIANGLGFSFVQMEKESCSHL
1	}						KSDLVRIKRLFPGHPAEENGAIAAGDIILGRE WEGPRKASSSRCRGSWAMQLSVOAGPSFAS
l	1						YYPAAVEVLHLLRGAPOEVTLLLCRPPPGAL
							PELEQEWQTPELSADKEFTRATCTDSCTSPIL
1							GSRGQLGGTVPPQMQGKAWGLRPESSQKAIR
							EGTMGAKTERDLGPVP
647	7	1997	A	4854	1044	335	PRVRGDWPLEKKKSNSNIHPIFSWCGSTDSKD

SEQ ID	SEQ ID	Met	SEQ	Predicted	Deadisted and	I Amino cold common (A A) in G G
NO: of	NO: of	hod	ID NO:	beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1 200	in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid,
cotide	seq-		USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence	1	09/496	correspondi	to last amino	I=Isoleucine, K=Lysine, L=Leucine,
uence	delice	Į	914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline,
delice		l	214	amino acid		Q=Glutamine, R=Arginine, S=Serine,
1		ļ		residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		Ì		peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
]			Ì			/=possible nucleotide deletion, \=possible
		<u> </u>		sequence	 	nucleotide insertion
					ì	IVMPTYDLTDSVLETMGRVSLDMMSVQANT
		j	1	[Ĭ	GPPWESKNSTAVWRGRDSRKERLELVKLSRK
ł		1	1			HPELIDAAFTNFFFFKHDENLYGPIVKHISFFD
1		1				FFKHKYQINIDGTVAAYRLPYLLVGDSVVLK
		1				QDSIYYEHFYNELQPWKHYIPVKSNLSDLLEK
1		ł	ł	}	ł	LKWAKDHDEEAKKIAKAGQEFARNNLMGD
ļ						DIFCYYFQTFPRNMPIYK
648	1998	Α	4867	2030	837	AGMLPAVGSADEEEDPAEEDCPELVPMETTQ
1		!				SEEEEKSGLGAKIPVTIITGYLGAGKTTLLNYI
	1	1				LTEQHSKRVAVILNEFGEGSALEKSLAVSQG
1	l .		ļ			GELYEEWLELRNGCLCCSVKDNGLRAIENLM
İ	ĺ	ĺ	1 1		ľ	QKKGKFDYILLETTGLADPGAVASMFWVDA
						ELGSDIYLDGIITIVDSKYGLKHLAEEKPDGLI
İ		!				NEATRQVALADAILINKTDLVPEEDVKKLRT
	1	i	Į į			TIRSINGLGQILETQRSRVDLSNVLDLHAFDSL
ľ	1	l	1 1			SGISLQKKLQHVPGTQPHLDQSIVTITFDVPG
		1	i i			NAKEEHLNMFIQNLLWEKNVRNKDNHCMEV
į		İ	i			IRLKGLVSIKDKSQQVIVQGVHELYDLEETPV
	.					CANDETERMINE AT LODAR DADAR NOT BY A
1 .		J				SWKDDTERTNRLVLLGRNLDKDILKQLFIAT
649	1999	Ā	4873	226	189	VTETEKQWTTHFKEDQVCT
045	1733	Λ.	40/3	220	109	DGVSLLLPKLGVQWAQYWAHWQPPLPGFKR
			!			FSCLSLRSSWD*KCAPPHPAFVFLVEMGFHRV
1						GQAGLELRTSGDPPASASQSAGITGVSHLA*P
650	2000	Ā	4874			TSMPLLPFQRLCVYI
000	2000	A	48/4	2	437	FFFLRRSFAFVAQAGVQWCDLGSPQPLPPGF
				. 1		K*FSCLSLPSSWDYRHAPPPCPS*FLYF**RQG
				, i		FTMLARLVLNS*PHDLPTSPSQSAEIKGVSHR
			li			CPASFYLFLKYYLEAKFCA*GECAPSAGVGA
L						GYKRGHKSCLLINCVVQI
651	2001	Α	4898	1701	771	DAWGPETRLARILNPDSFIEPRPGRLPELEATR
						PHMEPKASCPAAAPLMERKFHVLVGVTGSV
						AALKLPLLVSKLLDIPGLEVAVVTTERAKHFY
.						SPQDIPVTLYSDADEWEMWKSRSDPVLHIDL
						RRWADLLLVAPLDANTLGKVASGICDNLLTC
					J	VMRAWDRSKPLLFCPAMNTAMWEHPITAQQ
		1		1		VDQLKAFGYVEIPCVAKKLVCGDEGLGAMA
	Í					EVGTIVDKVKEVLFQHSGFQQS*PGISVMGVP
			i i			LYSEWVOAKSVKMDVGKIGGYPHLLNGGPA
L					İ	LSLPRGQACSRLNWTEGPGLSFFQPGEAAA
652	2002	A	4927	1	611	FRGRQTSRPARGFSPWRPPGTMQEPSSGECPA
	İ			- 1		SP*LPCASNRLAFGGLIFPCAPLVPYPAPFSPLL
				I		PAFSCAPRPRAHTHSRTHPSAPLVPKPSSRAR
! !	i	ļ		I		GQSPIPSRASSPSCSWAQVPGVALARCAGVC
1	·	- 1	.	1	ł	KPGDSWRVAACISGRCCSRGRRRGSGPRNPE
ľ	l	l				QSFRGAWGPSFWGSWKSQRELSAGGAQAWP
	ļ					LLGSAGSGLRGEA
653	2003	A	4965	2	283	
			7707	-	202	FFFFI*DGVSLCHPGWNAVARSWLTATSASR
	·	i	1		-	VQAVSCFRLPSSWDYRHATMPG*FF*YF**R
654	2004	Ā	4968		427	WGFTilailvlns*PQVicppwppkvltlqa
J. 7	-004	^	4700	3	437	RPGIPGRRFRRSWFCQLP*EPEPGLESLATPGD
į	[l	i		1	IPAVGLGALGVIPPVRVPQRPPTQRSQGRGW
ł	1	l	ì	ļ		DPERDPGCRVQVSRGPRFGEQKTPGLQGCLP
1	ŀ	ł	ŀ	l	1	PPCLTHLAAASCVVVWCGRWKRDSAECQCD
(55	2005	, 	400			HSCSAVSQQEDRCRSSSCS
655	2005	A	4983	201	397	MNNNTTCIQPSMISSMALPIIYILLCIVGVFGN
		لــــــــــــــــــــــــــــــــــــــ				TLSQWIFLTKIGKKTSTHIYLSHLVTANLLVC
656	2006	A	4988	332	159	LVHKDMYREFFEEEAQASNKHVTRCLTSLVI
					[REVHIKTMR*HFLPIRLEKNKNNIKD
657	2007	В	5008	129	465	MAGMKTASGDYIDSSWELRVFVGEEDPEAES

NO: of NO: of NO: of No: of N	SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
			•		f		D=Asnartic Acid F=Glutamic Acid
Sequence	N .			1 .			
uenice	cotide	seq-		USSN	location	corresponding	
maino acid residue of peptide residue of peptide sequence T-Threosine, V-Valine, W-Typosiban, Y-Stop codon, pepside sequence Threosine, V-Valine, W-Typosible nucleotide deletion, in-possible nucleotide deletion in-possible nucleotide deletion in-possible nucleotide deletion in-possible nucleotide deletion in-possible nucleotide deletion in-possible nucleotide deletion in-possible nucleot		uence	ł	1			M=Methionine, N=Asparagine, P=Proline,
Peptide Pept	uence			914			
Popide Sequence	1						T=Threonine, V=Valine, W=Tryptophan,
	ļ		1		1	sequence	Y=1 yrosine, X=Unknown, *=Stop codon,
			[1'			
Alwwegkrowllothwytolkyalladarlf FGPQHRPVILLIPMRAILIX*		 	1	 	Sequence		
658 2008 A 5017 1 292 FFFFKETESINSTYGAGYWHDIGSLQPPPPGF KRISCLSLLSSWDYRCAPPHPANFVELVETGF FFFKETESINSTYGAGYWHDIGSLQPPPPGF KRISCLSLLSSWDYRCAPPHPANFVELVETGF HPHVAGAGILXILT SANLGISTSIPPIFILS SANLGISTSIPPIFILS FFFKETESINSTYGAGYWHDIGSLQPPPPGF KRISCLSLLSSWDYRCAPPHPANFVELVETGF HPHVAGAGILXILT SANLGISTSIPPIFILS SANLGISTSIPPIFILS SANLGISTSIPPIFILS SANLGISTSIPPIFILS SANLGISTSIPPIFILS SANLGISTSIPPIFILS SANLGISTSIPPIFILS SANLGISTSIPPIFILS SANLGISTSIPPIFILS SANLGISTSIPPIFILS SANLGISTSIPPI				-	!		AIWWEOKROWLLOTHWTLDKYGILADARLE
658 2008 A 5017 1 292 FFFFKETESHSVTQAGV@WHDIGSLQPPPPAGE KRFSCSLLSSWDYRACPHPANFULFIGF HHVAQAGLKLLT.*SANLGLSTSLPPLFILES CREGGESLTGGTFGNWGDGLLVSEDWSILTF TYSLVSPULGKWSPCLQQFGLSAVHTWPWL MAACWAVHYKTHMRPGLAVLPRLVIANSWS AIILLWPPKALGLQ.		<u></u>		1	[Ĭ	FGPQHRPVILRLPNRRALRLX*
Son	658	2008	Α	5017	1	292	FFFFKETESHSVTQAGVQWHDLGSLOPPPPGF
659 2009 A 5018 17 338 RGHGGKSLTGGTPGNWGDGLLYSEDWSLILD RSVSPVLGKWSPCLQGPGLAXPHTWPWL MAACWAVHYKTHARPGLAYLPRLYLNSWS *AILLWPPKALGLQA MAACWAVHYKTHARPGLAYLPRLYLNSWS *AILLWPPKALGLQA MACWAVHYKTHARPGLAYLPRLYLNSWS *AILLWPPKALGLQA MACWAVHYKTHARPGLAYLPRLYLNSWS *AILLWPPKALGLQA MACWAVHYKTHARPGLAYLPRLYLNSWS AGSCRQKKTTPGTYVLCVCSFWYLQRGEPH HRTGARWNH RGSCSSTQAKYQWFHYGPLQSQPPGLKQSSQ LSLPNSRDHRHVPPRLAIFSFAETGSPYFAQAS LSLPNSRDHRHVPPRLAIFSFAETGSPYFAQAS LSLENSSAGARITGVSHRAWPLK*F NLNQYQTLTMN EINNGYFQMPLCNGGNLAVTGSWADRSPLH EAASQGRLIALRTLLSQGYNVANTLDHVTP LHEACLGDHVACARTLLEAGANVNAITIBGV TELNTACSQGSSSCAELLLEYGAQAQLESCLP SPTEEGASKGHHECLDILISWGIDVDQEPISWG TELNTACSQGSSSCAELLLEYGAQAQLESCLP SPTEEGASKGHHECLDILISWGIDVDQEPISWG TELNTACSQGSSSCAELLLEYGAQAQLESCLP SPTEEGASKGHHECLDILISWGIDVDQEPISWG TELNTACSQGSSCAELLLEYGAQAQLESCLP SPTEEGASKGHHECLDILISWGIDVDQEPISWG TELNTACSQGSSSCAELLLEYGAQAQLESCLP SPTEEGASKGHHECLDILISWGIDVDQEPISWG TELNTACSQGSSCAELLLEYGAQAQLESCLP SPTEEGASKGHHECLDILISWGIDVDQEPISWG TELNTACSQGSSSCAELLLEYGAQAQLESCLP SPTEEGASKGHHECLDILISWGIDVDQEPISWG TELNTACSQGSSCAELLLEYGAQAQLESCLP SPTEEGASKGHHECLDILISWGIDVDQEPISWG TELNTACSQGSSCAELLLEYGAQAQLESCLP SPTEEGASKGHHECLDILISWGIDVDQEPISWG TELNTACSQGSSCAELLLEYGAQAQLESCLP SPTEEGASKGHHECLDILISWGIDVDQEPISWG TELNTACSQGSSCAELLLEYGAQAQLESCLP SPTEEGASKGHHECLDILISWGIDVDQEPICH TELNTACSQGSTCAAVAVLSSILISWGIDVDQEPICH TELNTACSQGSTCAAVAVLSSILISWGIDVDQEPICH TELNTACSQGSTCAAVAVLSSILISWGIDVDQEPICH TELNTACSQGSTCAAVAVLSSILISWGIDVDQEPICH TELNTACSQGSTCAAVAVLSSILISWGIDVDQEPICH TELNTACSQGSTCAAVAVLSSILISWGIDVDQEPICH TELNTACSQGSTCAAVAVLSSILISWGIDVDQEPICH TELNTACSQGSTCAAVAVLSSILISWGIDVDQAAVAVLSGGSTCAAVAVLSSILISWGIDVGGGSTCAAVAVLSSILISWGIDVGGGGASAVAVLSGGGAGAAVAVLSGGAAVAVLSGGGAGAAVAVLSGGGAGAAVAVLSGGGAGAAVAVLSGGAAVAVLSGGAGAAVAVLSGGGAGAAVAVDCAAVAVLSGGGAGAAVAVLSGGGAAVAVAVLSGGGAAVAVAVLSGGGAAVAVAVLSGGGAAVAVAVLSGGGAAVAVAVLSGGGAAVAVAVLSGGGAAVAVAVACGGGAAVAVAVLSGGGAAVAVAVACGGAAVAVAVACGGAAVAVACGGAAVAVACGGAAVAVACGGAAVAVACGAAVAVACG							KRFSCLSLLSSWDYRCAPPHPANFVFLVETGF
PNSLVSPYLOKWSPCLQGPGLSAVHTWPWL MAACWAVHYKTHMRPOLAVLPRLVLNSWS AIIILWPPKALGLQA	650	2000	1.	6010		220	HHVAQAGLKLLTL*SANLGLSTSLPIPLFILLS
MAACWAYHYKTHARPOLAVLPRLYLNSWS	659	2009	A	2018	17	338	RGHGGKSLTGGTPGNWGDGLLVSEDWSHLIF
*AIILLWPPKALGLQA	1						1*NSLVSPVLGKWSPCLQGPGLSAVHTWPWL
2010 A 5028 2 310	1		ł			}	
HPGHLCLQPPGGPA*FLDVCRGCCPHPVPGST	660	2010	A	5028	2	310	
AGSCPRQKKTTPGPTVLCVCSFWIYQRGEPH HRTGARWH HRTGARWH A	}					***	HPGHLCLOPPGGPA*FLDYCRGCCPHPVPGST
HRTGARWNH	ļ		}				AGSCPROKKTTPGPTVLCVCSFWIYORGEPH
SILPINSPORRHYPPRIA AIFSFAETICSPYFAQAS LELLGSSHPYTSASQSARITGVSHRAWPLK+P	L		<u> </u>	·			HRTGARWNH
LSLPNSRDHRHYPPRLAIFSFAETGSPYFAQAS	661	2011	Α	5050	752	431	RQSCSSTQAKVQWFHYGPLQSQPPGLKQSSQ
NINQYQTLTMN	}						
A 5054 48 103			l				
EAASQGRLLARTILLSQGYNVNAVIDHVTP	662	2012	_	5054	40	102	
LHEACLGDHVACARTLLEAGANVNAITIDGY TPLFNACSQGSPSCAELLEYGAQAQESCLP SPTHEGASKGHHECLDILISWGIDVDQEIPHSG TPLYVACMAQQFHCIWNLIYAGAGVRKGKY WDTPLPGAGHQSTQKLe*LFAMVEIWQ WDTPLPGAGHQSTQKLe*LFAMVEIWQ WDTPLPGAGHQSTQKLe*LFAMVEIWQ WDTPLPGAGHQSTQKLe*LFAMVEIWQ WDTPLPGAGHQSTQKLe*LFAMVEIWQ WRS*SFAHCASVYKHHYMDGQTPCLFVSSK ADLPEGVAVSGPSPAEFCRKHRLPAPVPPSCA GPAFPSTHFTQLATMAAFPHLVHAELHPSSF WLRGLLGVVGAAVAVLSFSLYRVLVKSQ WLRGLLGVVGAAVAVLSFSLYRVLVKSQ GPAFPSTHFTQLATMAAFPHLVHAELHPSSF WLRGLLGVVGAAVAVAVSTSLYRVLVKSQ QLFVFILLIYLFILTGINAIIISTIVLDRALHTP MYFFLAILSCSEICYTFVIVFKMLVDLLSQKK TISFLGCAIQMFSFLFGSSISFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAC GFTVSLVTTSLVFHLPFHSSNQHE GFTVSLVTTSLVFHLPFHSSNQHE QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG PLORGLPSFNISLESVQLDSE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIALSEATAADVEAYDDCQPT KRTFATFGKDVSDHPAEFRPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC GFTVSVSYEIVDPGITARC GFTVSVSYEIVDPGITARC GFTVSVSYEIVDPGITARC GFTVSVSYEIVDPGITARC GFTVSVSYE	002	2012	^	3034	40	103	
TPLFNACSQGSPSCAELLLEYGAQAQLESCLP SPTHEBGASKGHHECLDILISWGIDVDQEIPHSG SPTHEBGASKGHHECLDILISWGIDVDQEIPHSG TPLYVACMAQQFHCIWNLIYAGAGVRKGKY WDTPLPGAGHQSTQKLE*LFAMYEIWQ VRNS*SPAHCASVFHIYMDGGTPCLFVSSK ADLPEGVAVSGPSPAEFCRKHRLPAPVFPSCA GPAEPSTTIFTQLATMAAPFHLVHAELHPSSF WLGGLIGVVGAAVAFFHLVHAELHPSSF WLGGLIGVVGAAVAFFHLVHAELHPSSF WLGGLIGVVGAVAVSFSLYRVLVKSQ QLLFVIFLLLYLFTLGTNAIHISTIVLDRALHTP MYFFLAILSCSEICYTFVIVFKMLVDLSQKK TISFLGCAJQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAC GFTVSLVTTSLVFHLPPHSSNQHE HDVQRSLYTTSLVFHLPPHSSNQHE HDVQRSLYGDTAAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIALSEATAADVEAYDDCGPT KRTFATFGKDVSDHFAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC							I HEACT COHVACADELL FACANIVIA PEDCY
SPTHEGASKGHHECLDILISWGIDVDÖEIPHSG	ł		i				TPLENACSOGSPSCAFLLIFYGAGAGI ESCLE
TFLYVACMAQQFHCIWNLIYAGAGVRKGKY WDTPLPGAGHQSTQKLE*LFAMVEIWQ				i 'i			SPTHEGASKGHHECLDILISWGIDVDOFIPHSG
WDTPLPGAGHQSTQKLE*IFAMYEIWQ			Ī				TPLYVACMAQQFHCIWNLIYAGAGVRKGKY
ADLPEGVAVSGPSPAEFCRKHRLPAPVPFSCA GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF WLRGLLGVVAAVAAVLSFSLYRVLVKSQ GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF WLRGLLGVVAAVAAVLSFSLYRVLVKSQ CLFVIFLLLYLFTLGTNAIIISTIVLDRALHTP MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAC GFTVSLVTTSLVFHLLPFHSSNQHE GFTVSLVTTSLVFHLPFHSSNQHE GFTVSLVTTSLVFHLPFHSSNQHE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNITIALSIEATAADVEAYDDCGPT KRTFATTGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPYTSPLHLKSSL PIKPSVSYEIVDPGITARRC GFTVSLVFHYDFGTTARC GFTVSLVFHYDFTTARC G							WDTPLPGAGHQSTQKLE*LFAMVEIWQ
GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF WRGLIGVVGAAVAAVLSTSLYRUVKSQ	663	2013	Α	5066	951	580	VRNS*SFAHCASVYKHHYMDGQTPCLFVSSK
WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ				1			ADLPEGVAVSGPSPAEFCRKHRLPAPVPFSCA
1			İ]]	. ,		GPAEPSTTIFTQLATMAAFPHLVHAELHPSSF
Coleytellylyteticgtnailistivldralhtp Coleytellylyteticgtnailistivldralhtp Myfflailscseicytfvivpkmlvdllsqkk Tisflgcaiqmfsflffgsshsfllaamgydr Ymaicnplrysvlmghgvcmglmaaawac Gftyslvttslvfilpfissnqhe Gftyslvttslvfilpfissnqhe Gftyslvttslvfilpfissnqhe Gftyslvttslvfilpfissnqhe Gftyslvttslvfilpfissnqhe Gftyslvttslvfilpfissnqhe Gftyslvttslvfilpfissnqhe Gftyslvttslvfilpfissnqhe Gftyslvttslvfilpfissnqhe Gftyslvttslvfilpfissnqhe Gftyslvttslvfilpfissnghe Gftyslvttslvfilpfissnghe Gftyslvttslvfilpfissnghe Gftyslvttslvfilpfissnghe Gftyslvttslvfilpfissnghe Gftyslvttslvfilpfissnghe Gftyslvttslvfilpfissngh Gf	664	2014		5071	550	1	WLRGLLGVVGAAVAAVLSFSLYRVLVKSQ
MYFFLAILSCSEICYTFVIVPKMLVDLLSQKK TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAC GFTVSLVTTSLVFHLPFHSSNQHE GFTVSLVTTSLVFHLPFHSSNQHE 665 2015 A 5074 496 692 QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEFRPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL		2011		50/1.	330	•	
TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR YMAICNPLRYSVLMGHGVCMGLMAAAWAC GFTVSLVTHLPFHSSNQHE 665 2015 A 5074 496 692 QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG PL*IQRGLPSFNSLEGHSLKDSGHESVQLDSE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGFKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRFTRPPLLTLLAHVSFEPAGFSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							MYFFLAILSCSEICYTFVIVPKMI.VDLI.SOKK
496 692 2015 A 5074 496 692 QYHNTGSAGHHAHCQVGHSPHVHYPSGCG PL*IQRGLPSFNSLEGHSLKDSGHEESVOLDSE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVYRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPITFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLINLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSFEPAGFSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGCGGDGLHHASAACASLPP SPILPVLLPELGPL			•	[]			TISFLGCAIQMFSFLFFGSSHSFLLAAMGYDR
GFTVSLVTTSLVFHLPFHSSNQHE				1			YMAICNPLRYSVLMGHGVCMGLMAAAWAC
PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL		2017					GFTVSLVTTSLVFHLPFHSSNQHE
HDVQRSLYCDTAVNDVLNTSVTSMGSQMPD HDQNEGFHCREECRLLGHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL	000	2015	Α	5074	496	692	QQYHNTGSAGHHAHCQVGHSPHVHYPSGCG
HDQNEGFHCREECRILGHSDRCWMPRNPMPI RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL				l			PL*IQRGLPSFNSLEGHSLKDSGHEESVQLDSE
RSKSPEHVRNIIALSIEATAADVEAYDDCGPT KRTFATFGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC 666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGY*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							HDOVECTUCE FOR HOUSE ON THE STATE OF THE STA
KRTFATFGKDVSDHPAEERPTLKGKRTVDVT ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC	İ				j		RSK SPEHVRNII A I SIE ATA ADVIE AVDDOODE
ICSPKVNSVIREAGNGCEAISPVTSPLHLKSSL PTKPSVSYEIVDPGITARRC						ł	
666 2016 A 5080 408 248 IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGFVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL							
VGPKAQLSGLQLQPCLYKRR 667 2017 A 5081 129 247 DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ PCVY*RR 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL					I		
667 2017 A 5081 129 247 DLTNS:HFFLFDFQKTGPPLGGPKAQFSSLQLQ 668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL	666	2016	A	5080	408	248	IMLLSTSS*VYFQSSTKDSHFFLFDFQKTGPPL
668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGY*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL	((2						VGPKAQLSGLQLQPCLYKRR
668 2018 A 5086 852 233 NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL	667	2017	Α	5081	129	247	DLTNSHFFLFDFQKTGPPLGGPKAQFSSLQLQ
RALPTTFADIENLKYLLFTRDASQPFYLGHTV IFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL	669	2010		FNOC		222	PCVY*RR
iFGDLEYVTVEGGIVLSRELMKRLNRLLDNSE TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGY*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL	500	2016	A	3080	832	233	NIKSNDRWVQIKTAYKYFF*KNGDNYNWVF
TCADQSVIWKLSEDKQLAICLKYAGVHAENA EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL	İ				ľ	1	
EDYEGRDVFNTKPIAQLIEEALSNNPQQVVEG CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL		į			j	1	TCADOSVIVEI SEDVOI AIG EVACUTA DE LA CONTRE LA
CCSDMAITFNGLTPQKMEVMMYGLYRLRAF GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL	ļ	l			ł	l	
GHYFNDTLVFLPPVGSEND 669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL]	1	Ī			l	
669 2019 A 5101 1 329 PGRPTRPPLLTLLAHVSPEPAGPSCDSLAQPG ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL				1		-	
ASGV*VQHDSHPPLLCGSQCLSEPVPGSHGPP RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL	669	2019	Α	5101	1	329	
RGCQHEAAPCPRGPGSDGLHHASAACASLPP SPILPVLLPELGPL					•		
	1	1			. 1		RGCQHEAAPCPRGPGSDGLHHASAACASLPP
O/O 2020 A 3102 3 347 DAWGNRCAVGAAPRLIHLHLCCTPADPSRKP	670	2020		5100	!		
	070	2020	A	3102	3	347	DAWGNRCAVGAAPRLIHLHLCCTPADPSRKP

SEQ ID NO: of NO: of nucl- cotide seq- uence SEQ ID NO: of nucl- cotide seq- uence SEQ ID NO: of nucleotide	ine, ne,
nucleotide sequence in ucleotide location corresponding to last amino acid residue of peptide residue of peptide sequence sequence in ucleotide location corresponding to last amino acid residue of peptide sequence sequence sequence in ucleotide location corresponding to last amino acid residue of peptide sequence seq	ne,
cotide sequence USSN 09/496 USSN 09/496 Osation corresponding to last amino acid residue of peptide residue of peptide sequence Osation corresponding to last amino acid residue of peptide sequence Osation corresponding to last amino acid residue of peptide sequence I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Prolife	ne,
sequence uence 09/496 correspondi ng to first amino acid residue of peptide sequence peptide sequence of peptide sequence sequence of peptide sequence sequence of peptide sequence sequence of peptide sequence sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide sequence of peptide of peptide sequence of peptide of peptide sequence of peptide of pepti	-
uence 914 ng to first amino acid residue of peptide residue of peptide sequence 92Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptopha Y=Tyrosine, X=Unknown, *=Stop code /=possible nucleotide deletion, \=possib nucleotide insertion	-
amino acid residue of peptide residue of peptide sequence y=Tyrosine, X=Unknown, *=Stop code /=possible nucleotide deletion, \=possible nucleotide insertion	
residue of peptide /-possible nucleotide deletion, \-possible nucleotide insertion	
peptide /-possible nucleotide deletion, \-possible sequence nucleotide insertion	1,
peptide /-possible nucleotide deletion, \-possible sequence nucleotide insertion	n.
sequence nucleotide insertion	le.
DEL*NMNGRVDYLVTEEEINLTRG	OSCI CENT
VGGTDQQYVSNDSGIYVSRIKENG	NAVI DOD
LQEGDKILSVNGQDLKNLLHQDAV	DI EDALA
CVA VICE PROVINCE OF THE PROVI	DLFKNA
GYAVSLRVQHRLQVQNGPIGHRGE	
FMVLVPVFALTMVAAWAFMRYRC	
671 2021 A 5105 672 400 RDGREELCLQQEPTLPSRICSSAPLL	YFLFICPF
VLLLLLISLLCLYWKARKLSTLRS	VTRKEKA
LWVDLKEAGGVTTNRMED*EEDEG	N .
672 2022 A 5148 72 314 IIYFSYNIFLKITELLNDVERLKOALI	GLSOLT
YTSGNPTKRQSQLIDTLQHQVKSLE	OOLAVS
NQAHGALQEYVLAPCS	~~~~
673 2023 A 5152 210 335 REILCSRIGRLNIV*MSLFPNLTCRLI	JAIDIVIDA
NHFVEVT	ATT INTEN
	TAXOT
SIGSSQQFAQAQLHPSSSASFTQASN	
QTLQPIGVTHVPVGASFASNTVGVC	
VGISVPSQHLSNSSQISGSGQIQLIGS	FGNHPS
MMTINNLDGSQIILKGSGQQAPSNV	SGGLLV
HRQTPNGNSLFGNSSSSPVAQPVTV	PFNSTNF
QTSLPVHNIIIQRGLAPNSNKVPINIQ	
GQQNTYNVNNLGIQQHHVQQGISF	
GSVVGPHMSVNIVNQQNTRKPVTS	•
GSIVIHSPMGQPHAPQSQFLIPTSLSV	
HVQTINGQLLQTQPSQLISGQVASE	
NSSNMLRTNQPYTGPMLNNQNTAV	INVICA
NSSIMILATING TIGHMUNDINIA	HLVSGQ
TFAASGSPVIANHASPQLVGGQMPL	
VLHLSPGQSSVSQGRPGFATMPSVT	
• FPAVSSASTAHPSLGSAVQSGSSGS	
TQPNRTPVPVSVSHRLPVSSSKSTST	FSNTPGT
GTQQQFFCQAQKKCLNQTSPISAPK	
	GTAQPQ
QEKVVGSSPGHPAVQVESHSGGQK	
LTKGAFILQQLQRDQAHTVTPDKSH	
AVQRLLSYHVCQGSMPTEEDLRKV	
ATOLLKRTOAMLNKYRCLLLEDAN	
MVMIDRMFNQEERASLSRDKRLAL	
ADFCCSFKLDKAAHETQFGRSDQHO	
I DRAW ACONDA LYMOLUMIA AND	OEIT 100
LQPPAKAQGRDRAKTGVTEPMNHD	QrnLVP
NHIVVSAEGNISKKTECLGRALKFD	
YQSTSEEKASRREPLKASQCSPGPEC	
RSDHGTESKLSSILADSHLEMTCNN	
RNSPKNEVLHTDIMKGSGEPQPDLQ	
TFKNILELKKAGRQPQSDPTVSGSVI	ELDFPNF
SPMASQENCLEKFIPDHSEGVVETD	SILEAAV
NSILEC	
675 2025 A 5154 599 1880 LKKMEPFSCDTFVALPPATVDNRIJF	GKNSDR
LYDEVQEVVYFPAVVHDNLGERLK	
OVPETYAVVLSRPAWLWGAEMGA	
GNEAVWGREEVCDEEALLGMDLVI	
DTAEKALNVIVDLLEKYGQGGNCTI	
SYHNSFLIADRNEAWILETAGKYWA	
GVRNISNQLSITTKIAREHPDMRNYA	
WDGKKEFDFAAAYSYLDTAKMMT	SSGRYCE
GYKLLNKHKGNITFETMMEILRDKF	SGINME
GEFLTTASMVFILPODSSLPCIHFFTO	
SVFKPFIFVPHISQLLDTSSPTFELEDI	
HFKPDRRHPLYQKHQQALEVVNNN	
MLDNMRKLEKELFREMESILONKHI	
	~ v cMv
NLFPQCTKDEIQIYQSNLSVKVSS	

NO of No o	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
Decidide Decidide					,		Annio acid sequence (A-Alamie C-Cysteme,
Sequence			1			1	
Sequence	4		1				
Benoc	1		1				i=isoleucine, K=Lysine, L=Leucine,
amino acid residue of peptide residue of peptide residue of peptide sequence T-Timeonine, V-Valine, W-Tyopine, N-Sop codon, peptide sequence Sequenc		uence		1		1	
Presidue of peptide Pe	uence	1	1	914			
Peptide Sequence		1	1	ł	amino acid	of peptide	
Peptide Sequence		i	İ	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
Sequence	1		ļ	J	peptide]	/=possible nucleotide deletion. \=possible
FFFFSCLSLPSSWDYRRPPRRANELYF**PRRG FTLLARWVIS**PHDPPASASQSAGTIGVSIRA RT	676	2026	Δ	5155		306	EERI DDCI AI CDDDDCCI OWDNI CCI OADDDC
FTLLARMVSIS*PHDPPASASQSAGITGVSHRA	0.0	2020] ^] 2133	-	300	ETTEROL OF DOCUMENT DEPEND A TO STATE OF
RPT	1		l	į .	l		
	ł		ı]	i	1	
	L						
SNDLDVPVGHIVHTGMLNEGGGYENDCSLA NLLEDVTWKYTALNILGPRAVDVLSELSY NLLEDVTWKYTALNILGPRAVDVLSELSY NLLEDVTWKYTALNILGPRAVDVLSELSY NLLEDVTWKYTALNILGPRAVDVLSELSY NLLEDVTWKYTALNILGPRAVDVLSELSY MPTPPHPFSLFCKENSVGYANGRYMSMTHT GEPGPMLYPIEYRWGFIMLSTLVSNS FALCRILGROMIVCVADFGLSKKIYSGDTYRO GRIAKMPVKWIAGELGRYYTSKSDWAFG VTMWEIATRGMTPYPCVQNHEMYDVLLHG HRLKQPEDCLDELCKI**9PQSP HPROPEDCIBLECKI**9PQSP HPROPEDCIBLECKI**9PQSP SESQVKHFKMRKIDLCISSEGSEVILATSSD KHPPEPNIDGNPETFVTTIGMPPGFIICFHKH VRIERLVIQSYFVQTILKIEKSTSKEPVDFEIQWI EXDLVHTEGQLQHOEIVAHDGSATYLFTICH HRLKQPEDCLDELCKI**9PQSP KHPPEPNIDGNPETFVTTIGMPPGFIICFHKH VRIERLVIQSYFVQTILKIEKSTSKEPVDFEIQWI EXDLVHTEGQLQHOEIVAHDGSATYLFTICH HRLKQPEDCLDELCKI**9PQSP KHPPENIDGNPETFVTTIGMPPGFIICFHKH VRIERLVIQSYFVQTILKIEKSTSKEPVDFEIQWI EXDLVHTEGQLQHOEIVAHDGSATYLFTICH HILLISYFILSALIKAVTSAACKASTICSAH RTVVVVFYGTISFMYFKFKAKDRNVDKTVAL FIGVYVTSINPITYSIRNAEVKAAVITLIRGG LISKRASHCVCCPLPISAGIG SETTILSYFITAGGROWIT*GREF DISMAMMYEPCLYFK GES 2031 A 5210 Z 231 FFFETESYSTQAGVQWPNLSSLKTLPPGFK*F SCLSLPSSWDYRCLPPCPANFCIFSRNOVLPC SCLSLPSSWDYRCLPPCPANFCIFSPSPLSCSSI SCLSLPSSWDYRCHARACTORYLEGILAGATA SCLSLPSSWDYRCHARACTORYLEGILAGATA SCLSLPSSWDYRCHARACTORYLEGILAGA	677	2027	A	5167	97	740	FFHSVDLLALEQSKTFYKPDWFDIVESEVKCC
SNDLDVPVGHIVHTGMLNEGGGYENDCSLA NLLEDVTWKYTALNILGPRAVDVLSELSY NLLEDVTWKYTALNILGPRAVDVLSELSY NLLEDVTWKYTALNILGPRAVDVLSELSY NLLEDVTWKYTALNILGPRAVDVLSELSY NLLEDVTWKYTALNILGPRAVDVLSELSY MPTPPHPFSLFCKENSVGYANGRYMSMTHT GEPGPMLYPIEYRWGFIMLSTLVSNS FALCRILGROMIVCVADFGLSKKIYSGDTYRO GRIAKMPVKWIAGELGRYYTSKSDWAFG VTMWEIATRGMTPYPCVQNHEMYDVLLHG HRLKQPEDCLDELCKI**9PQSP HPROPEDCIBLECKI**9PQSP HPROPEDCIBLECKI**9PQSP SESQVKHFKMRKIDLCISSEGSEVILATSSD KHPPEPNIDGNPETFVTTIGMPPGFIICFHKH VRIERLVIQSYFVQTILKIEKSTSKEPVDFEIQWI EXDLVHTEGQLQHOEIVAHDGSATYLFTICH HRLKQPEDCLDELCKI**9PQSP KHPPEPNIDGNPETFVTTIGMPPGFIICFHKH VRIERLVIQSYFVQTILKIEKSTSKEPVDFEIQWI EXDLVHTEGQLQHOEIVAHDGSATYLFTICH HRLKQPEDCLDELCKI**9PQSP KHPPENIDGNPETFVTTIGMPPGFIICFHKH VRIERLVIQSYFVQTILKIEKSTSKEPVDFEIQWI EXDLVHTEGQLQHOEIVAHDGSATYLFTICH HILLISYFILSALIKAVTSAACKASTICSAH RTVVVVFYGTISFMYFKFKAKDRNVDKTVAL FIGVYVTSINPITYSIRNAEVKAAVITLIRGG LISKRASHCVCCPLPISAGIG SETTILSYFITAGGROWIT*GREF DISMAMMYEPCLYFK GES 2031 A 5210 Z 231 FFFETESYSTQAGVQWPNLSSLKTLPPGFK*F SCLSLPSSWDYRCLPPCPANFCIFSRNOVLPC SCLSLPSSWDYRCLPPCPANFCIFSPSPLSCSSI SCLSLPSSWDYRCHARACTORYLEGILAGATA SCLSLPSSWDYRCHARACTORYLEGILAGATA SCLSLPSSWDYRCHARACTORYLEGILAGA		l					KEAVCVIDMSSFTEFEITSTGDQALEVLQYLF
LNKRSFMSPTDQQVHCWAWLKKIIMPKDS	ł						SNDLDVPVGHIVHTGMLNEGGGYENDCSIAR
	ł				ł		
PMTPDIFPSLE/CKEMSY/ANGIRVMSMTHT GEPGMALTPPEYRWGFTMLSTLVSNS 678	1	J	j	1		}	
GEFGMLYIPEYRWGFILSTLYSINS		[l		1		
678 2028 A 5183 1919 2018 PALCRLRDMTVCVADFGLSKKIYSGDVYRQ GRIAKMPVKWIAISSLADRVYTSKSDVWAFG VTMWEIATRGMTPYFGVQNHEMYDYLLHG VTMWEIATRGMTPYFGVQNHEMYDYLLHG FOR THE WITTOM THE WITT		ŀ	i	Ì	ŀ		
GRIAKMPYKWIAIESLADRYYTSKSDVWAFG				<u> </u>			
	678	2028	Α	5183	1919	2018	
		1	ĺ			1	GRIAKMPVKWIAIESLADRVYTSKSDVWAFG
HRILKOPEDCLIDELCKI**SPQSP RESQVKHFKMRKIDLCLSSEGSEVILATSSDE KHPPENIIDGNPETFWTTIGMPQEFIICFHKH VRIERLVIQSYFVOTILKIEKSTSKEPVDFEQWT EKDLVHTEGQLQNEEIVAHDGSATYLRFIIVS AFDHFASVHSVSAEGTVVSNLSS EILAVILKLACQBISLNALALMVATAVLTLAPI LLICLSYLFILSALRVPSAAGRCKAFSTCSAH RTVVVVPYGTISFMYFKPKAKDPYNDKTVAL FYGVYTPSLINJYSLIRAAGRCKAFSTCSAH RTVVVVPYGTISFMYFKPKAKDPYNDKTVAL FYGVYTPSLINJYSLIRAAGRCKAFSTCSAH RTVVVVPYGTISFMYFKPKAKDPYNDKTVAL FYGVYTPSLINJYSLIRAAGRCKAFSTCSAH RTVVVVPYGTISFMYFKPKAKDPYNDKTVAL FYGVYTPSLINJYSLIRAAGRCKAFSTCSAH RTVVVVPYGTISFMYFKPKAKDPYNDKTVAL FYGVYTPSLITAJGRGRWIT*GREF DTSMANMVKPCLYRK FFETESVSTITAGROWPNLSSLKTLPPGFK*F SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS FFETESVSTITAGROWPNLSSLKTLPPGFK*F SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF MIILLRSMYRINVKPYFFI*LFFSRVNC*SVIIG YARCYTFLIF*LF*LF*LF*LF*LF*LF*LF*LF*LF*LF*LF*LF*LF		1	ł	ł	}	1	
			l			ļ	
KHPPENIIDGNPETFWTTTGMFPQEFIICFHKH	679	2029	A	5190	39	499	
	""	2027	١.,	-1.70	, ,	***	
			Ì				
AFDHFASVHSVSAEGTVVSNLSS	Į.			i			
680 2030 A 5204 541 92 EILAVLKI_ACGDISLNALIALMYATAVLTLAPIL LICLSYJFILSAILRVPSAAGRCKAFSTCSAH RTVVVYFYGTISPMYKPKRAKDPNVDKTVAL FYGVVTPSLNPIIYSLRNAEVKAAVLTLLRGG LLSRKASHCYCCPLPLSAGIG 681 2031 A 5207 10 247 VPDNGDVTKLPYCSTLVEETSLTVSEAMEQSI KNESPLPGTLAHTCNTSTLGGGGRWIT*GREF DTSMAMMVKPCLYPK 682 2032 A 5210 Z 231 FFFETESYSITQAGVQWPNLSSLKTLPPGFK*F SCLSLPSSWDVRCLPPCPANFCIFSRNGVLPC WFGWSRTPDLS 683 2033 A 5218 85 402 CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF MIILLRSMYRRNVKPYFFI*LFSRVNC*SVIIG YARCYTFLIF*LFL*IPADSPITQEPKTVMLSK QSESAI 684 2034 A 5220 1 194 NLMKEMQNLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS 685 2035 A 5228 260 440 LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPNTTELSKCS*S*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAKLSSFRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSUMLHIRR KHSPITVTVWERELRKAKPRTKLITFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSWERSK GYFAKLSTELYI PRKREKKHVSLSWERSK GYFAKLSVERSK GYFAKLSVERSK QKMQQFQF SGCSKHVKQMKARQMKLSNTGEYESQRFR ASSQSAPSPDVGSQVT A 5244 1 428 MAAVVAATALKGRGARNARVL		}	ľ	l			EKDLVHTEGQLQNEEIVAHDGSATYLRFIIVS
LLICLSYLFILSAILRVPSAAGRCKAFSTCSAH RTVVVVFYGTISFMYFKPKAKDPNVDKTVAL FYGVVVFYGTISFMYFKPKAKDPNVDKTVAL FYGVVVFYGTISFMYFKPKAKDPNVDKTVAL FYGVVTPSLNPIPISLRNAEVKAAVLTLLRGG LLSRKASHCYCCPLPLSAGIG LLSRKASHCYCCPLPLSAGIG LLSRKASHCYCCPLPLSAGIG LLSRKASHCYCCPLPLSAGIG LLSRKASHCYCCPLPLSAGIG LLSRKASHCYCCPLPLSAGIG VPDNGDVTKLPVCSTLVEETSLTVSEAMEQSI KNESPLPGTLAHTINSTLGGRGRWIT*GFEF DTSMANMVKPCLYRK ENSEMPTICAGROWPWINSSLKTLPPGFK*F SCLSLPSSWDVRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS CSLSLPSSWDVRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPLS WPGWSRTPDLS WPGWSRTPDLS WPGWSRTPLS WPGWSRTPLS WPGWSRTPDLS WPGWSRTPLS WPGWSRTPLS WPGWSRTPLS WPGWSRTPLS WPGWSRTPLS WPGWSRTPLS WPGWSRTPLS WPGWSRTPLS WPGWSRTPLS WPGWSRTPLS WPGWSRTPLS WPGWSRTPLS WPGWSRTPS WPGWSTPS WPGWSTPS WPGWSTPS WPGWSTPS WPGWSTPS WPGWSTPS WPGWSTPS WPGWSTPS WPGWSTPS WPGWSTPS WPGWSTPS WP		<u></u>		<u></u>	<u> </u>		AFDHFASVHSVSAEGTVVSNLSS
LLICLSVLFILSAILRVPSAAGRCKAFSTCSAH	680	2030	Α	5204	541	92	EILAVLKLACGDISLNALALMVATAVLTLAPL
RTVVVVPYGTISFMYFKPKAKDPNVDKTVAL FYGVVTPSLNPIPYSLNAEVKAAVLTLLRGG LLSRKASHCYCCPLPLSAGIG LLSRKASHCYCCPLPLSAGIG LLSRKASHCYCCPLPLSAGIG VPDNGDVTKLPVCSTLVEETSLTVSEAMEQSI KNESPLPGTLAHTCTSLGGGRRWIT*GREF DTSMANMVKPCLYRK SKNESPLPGTLAHTCTSLGGGRRWIT*GREF DTSMANMVKPCLYRK SKNESPLPGTLAHTCTSLGGGRRWIT*GREF DTSMANMVKPCLYRK SKNESPLPGTLAHTCTSLGGGRRWIT*GREF DTSMANMVKPCLYRK SKNESPLPGTLAHTCTSLGGGRRWIT*GREF DTSMANMVKPCLYRK SKNESPLPGTLAHTCTSLGGGRRWIT*GREF DTSMANMVKPCLYRK SKNESPLPGTLAHTCTSLGGGRRWIT*GREF DTSMANMVKPCLYRK SKNESPLPGTLAHTCTSLGGGRRWIT*GREF SKLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS SCLSLPSSWDYRINDVEXAVSNIF MILLKSMYRINVKPYFFI*LFFSRVNC*SVIIG YARCYTFLIF*LFL*BFSRVNC*SVIIG YARCYTFLIF*LFL*BFSRVNC*SVIIG YARCYTFLIF*LFL*BFSRVNC*SVIIG YARCYTFLIF*LFL*BFSRVNC*SVIIG YARCYTFLIF*LFL*BFSRVNC*SVIIG YARCYTFLIF*LFL*BFSRVNC*SVIIG YARCYTFLIF*LFL*BFSRVNC*SVIIG YARCYTFLIF*LFL*BFSRVNC*SVIIG YARCYTFLIF*LFL*BFSRVNC*SVIIG YARCYTFLIF*LFL*BFSRVNC*SVIIG YARCYTFLIF*LFL*BFSRVNC*SVIIG YARCYTFLIF*LFT*LFT*LFT*LSKC*S*S*KINSGERE SCSSS SKANGGRAF YARCYTFLIF*LFT*LSKC*S*S*KINSGERE SCSSS SKANGGRAF YARCYTFLIF*LFT*LSKC*S*S*KINSGERE SCSSS SKANGGRAF YARCYTFLIF*LFT*LSKC*S*S*KINSGERE SCSSS SKANGGRAF YARCYTFLIF*LFT*LSKC*S*SEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW ORKJQVLADQKEGVSGCVTPTESLASLCTTQS ETTDLSSASCLRGFPMEKLQIVKPLEGSQTLY HWQQLAQPNLGTLDPRPGVLTKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPTTSAGGPVTVATANPGKCLSCT NSTFTFTTCRLHFSDITQVTPSSGFPSLSCGSS STVTGIFLPTTSAGGPVTVATANPGKCLSCT NSTFTFTTCRLHFSDITQVTPSSGFFSLSCGSS			ł				
FYGVVTPSLNPIIVSLRNAEVKAAVLTLLRGG LLSRKASHCYCCPLPLSAGIG			1				
LLSRKASHCYCCPLPLSAGIG			[l .			
681 2031 A 5207 10 247 VPDNGDVTKLPVCSTLVEETSLTVSEAMEQSI KNESPLPGTLAHTCHTSTLGGRGRWIT*GREF DTSMANMVKPCLYRK 682 2032 A 5210 Z 231 FFFETESYSITQAGVQWPNLSSLKTLPPGFK*F SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS 683 2033 A 5218 85 402 CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF MILLRSMYRINVKPYFFI*LFFSRVNC*SVIIG YARCYTFLIF*LF*IPADSPTDQEPKTVMLSK QSESAI 684 2034 A 5220 1 194 NLMKEMQNLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS 685 2035 A 5228 260 440 LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPNFTELSKCS*S*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHIR KHSRPIVTVWERELRKAKPNRKLTFLVLAND VIQNSKRKGPFTKDFAPVIVEAFKHVSSETD ESCKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRAI QSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKAQNMRLSNTGEYESQRFR ASSQSAPSPDVGSOVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSLATALHRLSLRRQNYLSEKQFFAEEW QKKIQVLADQKEGVSGCVTT						1	
KNESPLPGTLAHTCNTSTLGGRGRWIT*GREF DTSMANMVKPCLYRK	601	2021		5207	10	242	
DTSMANMVKPCLYRK	001	2031	Α.	3207	10 ,	247	
682 2032 A 5210 Z 231 FFFETESYSITQAGVQWPNLSSLKTLPPGFK*F SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTPDLS 683 2033 A 5218 85 402 CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF MIILLRSMYRINVKPYFFI*LFFSRVNC*SVIIG YARCYTFLIF*2LFL*IPADSPTDQEPKTVMLSK QSESAI 684 2034 A 5220 I 194 NLMKEMQNLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS 685 2035 A 5228 260 440 LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINFTELSKCS*S*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 I 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGPFCGILLFILTKR EVDKDRVKQMKARQNIMLSNTGEYESQRFR ASSQSARSPDVCSGVQT 688 2038 A 5249 I 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS ETTDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQLAQPNLGTILDPRPOVITKGFTQLPGD AIYHISDLEEDEEGTITFQVQQPLEVEEKLSTS KVPTGILIPPITSAGGPVTVATANPGKCLSCT NSTFTTCRILHPSDITQVTPSSGFPSLSCGSS							
SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTDLS							
SCLSLPSSWDYRCLPPCPANFCIFSRNGVLPC WPGWSRTDLS	682	2032	Α	5210	2	231	FFFETESYSITQAGVQWPNLSSLKTLPPGFK*F
WPGWSRTPDLS	j j						
683 2033 A 5218 85 402 CPSVSGLIKSDLRRHNINIGITNVDVKAVSNIF MIILLRSMYRINVKPYFF1*LFFSRVNC*SVIIG YARCYTFLIF*LFTL*IPADSPTDQEPKTVMLSK QSESAI 684 2034 A 5220 1 194 NLMKEMQNLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS 685 2035 A 5228 260 440 LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINTTELSKCS*5*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNTRALQSHSSPEGKEEPEPLSPLEFLY PRKAGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRQNYLSEKQFFAEEW QRLQVLADQKEGVSGCVIPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEDEEGITFQVQQPLEVEEKLSTS KPYTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTTTCRILHPSDITQVTPSSGFPSLSCGSS							
MIILLRSMYRINVKPYFF1*LFFSRVNC*SVIIG	683	2033	Α	5218	85	402	
	"	2000		5210	03	402	
QSESAI							
684 2034 A 5220 1 194 NLMKEMQNLNSENHKTWEEYKDTK*IMSYF YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS 685 2035 A 5228 260 440 LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINFTELSKCS*S*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVIPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITTQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	! [
YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL TDS 885 2035 A 5228 260 440 LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINFTELSKCS*5*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPELSPLEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVIPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGTIFQVQQPLEVEKKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	-						
685 2035 A 5228 260 440 LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINFTELSKCS*S*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAAKLSSPRPHRVGRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	084	2034	A	5220	1	194	
685 2035 A 5228 260 440 LHSQDGNSDPRKPQGEMSAHAFPVQTCGEED QKKTPQVPINFTELSKCS*S*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPELSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS					İ		YG*ALNVIKMAVLPKLMYRFSATLVKIPQHL
QKKTPQVPINFTELSKCS*S*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS							TDS
QKKTPQVPINFTELSKCS*S*KIMSGERE 686 2036 A 5239 79 508 GGEAAARAAKLSSPRPHRVGRRERGVGGMS AFSEAALEKKLSELSNSQQSVQTLSLWLIHHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	685	2035	Α	5228	260	440	LHSODGNSDPRKPOGEMSAHAFPVOTCGEFD
686 2036 A 5239 79 508 GGEAAARAAKLSSPRPHRVGRRERGVGGMS	[.			_			
AFSEAALEKKLSELSNSQQSVQTLSLWLIFHR KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRAI.QSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	686	2036	A	5230	79	508	
KHSRPIVTVWERELRKAKPNRKLTFLYLAND VIQNSKRKGPEFTKDFAPVIVEAFKHVSSETD ESCKKHLGRVLSIWEERS 687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRALQSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	°°° .	2000	**	2237	''	200	
687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRAI QSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLA WAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVIPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKLGESQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS							
687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRAI.QSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEVAGSSQKMGQPGP SGDSLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS							
687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRAI.QSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS							
687 2037 A 5244 1 428 MAAVVAATALKGRGARNARVLRGILAGATA NKASHNRTRAI.QSHSSPEGKEEPEPLSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS							ESCKKHLGRVLSIWEERS
NKASHNRTRALQSHSSPEGKEEPELSPELEYI PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	687	2037	A	5244	1	428	
PRKRGKNPMKAVGLAWAIGFPCGILLFILTKR EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS				· /	í	l	
EVDKDRVKQMKARQNMRLSNTGEYESQRFR ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS					j		
ASSQSAPSPDVGSGVQT 688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQFLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS					ļ	ŀ	· · · · · · · · · · · · · · · · · · ·
688 2038 A 5249 1 1407 LQQTEDKSLLNQGSSSEEVAGSSQKMGQPGP SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS							
SGDSDLATALHRLSLRRQNYLSEKQFFAEEW QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	(00	2020		5046			
QRKIQVLADQKEGVSGCVTPTESLASLCTTQS EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS	880	2038	A	5249	1	1407	
EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS					l	.	
EITDLSSASCLRGFMPEKLQIVKPLEGSQTLY HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS					l		QRKIQVLADQKEGVSGCVTPTESLASLCTTOS
HWQQLAQPNLGTILDPRPGVITKGFTQLPGD AIYHISDLEEDEEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS					l		
AIYHISDLEEDEEGITFQVQQPLEVEEKLSTS KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS					l		
KPVTGIFLPPITSAGGPVTVATANPGKCLSCT NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS				ŧ	l		
NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS NSTFTFTTCRILHPSDITQVTPSSGFPSLSCGSS					Į.	Ì	
					j	.	1
GSSSSNTAVNSPALAYRLSIGESITNRRDSTTT	•						
							GSSSSNTAVNSPALAYRLSIGESITNRRDSTTT

SEQ II	O SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of		hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan
1	ĺ	Ĭ	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
ł	- 1	1	Ì	peptide	1	/=possible nucleotide deletion, \=possible
		<u> </u>		sequence		nucleotide insertion
1		1	1			FSSTMSLAKLLQERGISAKVYHSPISENPLQPL
		l		1	1	PKSLAIPSTPPNSPSHSPCPSPLPFEPRVHLSEN
İ		1				FLASRPAETFLQEMYGLRPSRNPPDVGOLKM
]	1	ļ	,	NLVDRLKRLGIARVVKNPGAQENGRCOEAEI
						GPQKPDSAVYLNSGSSLLGGLRRNQSLPVIM
689	2039	+	5054	<u> </u>		GSFAAPVCTSSPKMGVLKED
009	2039	Α	5254	2	2621	LSLFGSRALGRSGARAMAKAKKVGARRKAS
	1	1				GAPAGARGGPAKANSNPFEVKVNRQKFQILG
	l	1				RKTRHDVGLPGVSRARALRKRTQTLLKEYKE
1	1	ĺ	1		ĺ	RDKSNVFRDKRFGEYNSNMSPEEKMMKRFA
1	}	i	1 1		}	LEQQRHHEKKSIYNLNEDEELTHYGQSLADIE
1					1	KHNDIVDSDSDAEDRGTLSGELTAAHFGGGG
1					1	GLLHKKTQQEGEEREKPKSRKELIEELIAKSK
1	1	!				QEKRERQAQREDALELTEKLDQDWKEIQTLL
1	1		1 1			SHKTPKSENRDKKEKPKPDAYDMMVRELGF
		i	}		1	EMKAQPSNRMKTEAELAKEEQEHLRKLEAE
			i			RLRRMLGKDEDENVKKPKHMSADDLNDGFV
			1 1		ĺ	LDKDDRRLLSYKDGKMNVEEDVQEEQSKEA
1	1)				SDPESNEEEGDSSGGEDTEESDSPDSHLDLES
1	1					NVESEEENEKPAKEQRQTPGKGLISGKERAG
1			1 !			KATRDELPYTFAAPESYEELRSLLLGRSMEEQ
1		1	1			LLVVERIQKCNHPSLAEGNKAKLEKLFGFLLE
1		}				YVGDLATDDPPDLTVIDKLVVHLYHLCQMFP
	ļ		1			ESASDAIKFVLRDAMHEMEEMIETKGRAALP
			[1			GLDVLIYLKITGLLFPTSDFWHPVVTPALVCL
1			1 1			SQLLTKCPILSLQDVVKGLFVCCLFLEYVALS
ļ] }	<i>'</i>		QRFIPELINFLLGILYIATPNKASQGSTLVHPFR
			! !			ALGKNSELLVVSAREDVATWQQSSLSLRWA
İ	1 1		1 1	•		SRLRAPTSTEANHIRLSCLAVGLALLKRCVLM
Į.					·	YGSLPSFHAIMGPLRALLTDHLADCSHPQELQ
						ELCQSTLTEMESQKQLCRPLTCEKSKPVPLKL
			1 1			FTPRLVKVLEFGRKQGSSKEEQERKRLIHKHK
· .	1 1		1			REFKGAVREIRKDNOFLARMOLSEIMERDAE
690	2040	A	5261	1	304	RKRKVKQLFNSLATQEGEWKALKRKKFKK
			5201	•	304	FFFFVFLVETGFHHVGQAGLELLTSGDPPTW
	1 1		i			ASQSAGITGVSHCSWPVIYVLSTLLHAVRNVL FKRTFPLKSSSFLSYDKEIFPILIVLKFYLVTLT
	1 1		1 1	i	i	SFVK
691	2041	A	5270	3	158	NCHTTHCTANWVHLPGTPPGWKIDGPAAAL
1 '		-		-		EVLSSFFFFFLKFSYKPQNIV
692	2042	Ā	5282	56	1268	GMEPVGCCGECRGSSVDPRSTFVLSNLAEVV
I			[]		-300	ERVLTFLPAKALLRVACVCRLWRECVRRVLR
ł	1			l	1	THRSVTWISAGLAEAGHLEGHCLVRVVAEEL
!]			l	1	ENVRILPHTVLYMADSETFISLEECRGHKRAR
1	1		[KRTSMETALALEKLFPKQCQVLGIVTPGIVVT
				İ	i	PMGSGSNRPQEIEIGESGFALLFPQIEGIKIQPF
i	1			}	ļ	HFIKDPKNLTLERHQLTEVGLLDNPELRVVLV
]]			ŀ	ļ	FGYNCCKVGASNYLQQVVSTFSDMNIILAGG
	[ĺ	Ī	QVDNLSSLTSEKNPLDIDASGVVGLSFSGHRI
	1			ļ		QSATVLLNEDVSDEKTAEAAMQRLKAANIPE
	1			}		HNTIGFMFACVGRGFQYYRAKGNVEADAFR
1						KFFPSVPLFGFFGNGEIGCDRIVTGNFILRKCN
l				l	ľ	EVKDDDLFHSYTTIMALIHLGSSK
693	2043	Α	5301	362	507	EEIKERFGPGLVIYWYGFIQELDCNRERGILLK
	1				- • ·	ACFPTNIVTLCHSIA
694	2044	A	5310	1	204	RVLTAINHTLKENLRKFYKGKKDKPLDLRPK
	[ĺ		1		KTRAMRRRLNMHEENLKTKKQHRKERLYPL
		ł	1	ł	Į.	RKYAAKA
695	2045	Α	5315	125		ETRSTAVKSEVQVCISLLLCLEDRTMPKKAKP
	·					

SEQ ID	SEO ID	Met	SEQ	Predicted	Deadine	I Amino cold compress (A. Alester C. C.
NO: of	NO: of	hod	ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
		1100	ŧ.			D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
i		1	ł	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	1	j	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1			peptide	1	/=possible nucleotide deletion, \=possible
1	1	1	ļ	sequence	1	nucleotide insertion
	+		 			TGSGKEEGPAPCKQMKLEAAGGPSALNFDSP
1	1		l	1	ľ	SSLFESLISPIKTETFFKEFWEQKPLLIQRDDPA
	j		ĺ			I ATVICOLTULTU KULOODOL GULOODIA
	Į.	1				LATYYGSLFKLTDLKSLCSRGMYYGRDVNV
1	i			Į		CRCVNGKKKVLNKDGKAHFLQLRKDFDQKR
		İ		ĺ		ATIQFHQPQRFKDELWRIQEKLECYFGSLVGS
			i			NVYITPAGSQGLPPHYDDVEVFILQLEGEKH
!	1	1	1	}	1	WRLYHPTVPLAREYSVEAEERIGRPVHEFML
1	l l]				KPGDLLYFPRGTIHQADTPAGLAHSTHVTIST
1	1	1				YQNNSWGDFLLDTISGLVFDTAKEDVELRTG
I	1	ŀ				IPRQLLLQVESTTVATRRLSGFLRTLADRLEG
İ	ĺ					TKELLSSDMKKDFIMHRLPPYSAGDGAELSTP
	1					GGKLPRLDSVVRLQFKDHIVLTVLPDQDQSD
ł	1		ł			ETQEKMVYIYHSLKNSRETHMMGNEEETEFH
	[İ			GLRFPLSHLDALKQIWNSPAISVKDLKLTTDE
	}	l			i	EKESLVLSLWTECLIQVV
696	2046	A	5318	1476	742	
030	2040	^	2210	14/0	/42	LMKXYLEAAELGEISDIHTKLLRLSSSQGTIET
	1				}	SLQDIDSRLSPGGSLADAWAHQEGTHPKDRN
1	ļ	ļ				VEKLQVLLNCMTEIYYQFKKDKAERRLAYN
	İ	ļ				EEQIHKFDKQKLYYHATKAMTHFTDECVKK
						YEAFLNKSEEWIRKMLHLRKQLLSLTNQCFDI
l						EEEVSKYQEYTNELQETLPQKMFTASSGIKHT
ł	1					MTPIYPSSNTLVEMTLGMKKLKEEMEGVVKE
						LAENNHILESGGSLTMDGGLRNVDCL
697	2047	A	5320	244	478	LDYNFFLFEMTFGLVSQAGVQWHDLGSLOPP
i				•		PPGFKQFSCLSLPSSWDYRHLPPHLANFSREG
1						VSPSWPGWSRTPDFR
698	2048	A	5324	266	714	LPIRKSLRSVRSGFPTSQSPITRNLDGTASGSC
	1				'''	LAKTVTGSLFRINVGLRGLVAGGIIGALLGTP
.}						VGGLLMAFQKYSGETVQERKQKDRKALHEL
ļ	1					KLEEWKGRLQVTEHLPEKIESSLQEDEPENDA
1						
699	2049	A	5334	699	277	KKIEALLNLPRNPSVIDKQDKD
033	2049	A	3334	699	211	RPHGHLVCISSSAGLSGVNGLADYCASKFAA
	1					FGFAESVFVETFVQKQKGIKTTIVCPFFIKTGM
						FEGCTTGCPSLLPILEPKYAVEKIVEAILQEKM
						YLYMPKLLYFMMFLKSFLPLKTGLLIADYLGI
						LHAMDGFADQKK
700	2050	Α	5344	3	614	PTAEEMSSLTPESSPELAKRSWFGNFISLDKEE
						QIFLVLKDKPLSSIKADIVHAFLSIPSLSHSVLS
1				l	İ	QTSFRAEYKASGGPSVFQKPVRFQVDISSSEG
				ļ		PEPSPRRDGSGGGGIYSVTFTLISGPSRRFKRV
1	1			ļ	<u> </u>	VETIQAQLLSTHDQPSVQALADEKNGAQTRP
1	1			İ		AGAPPRSLQPPPGRPDPELSSSPRRGPPKDKK
1				,	l	LLATNGTPL
701	2051	A	5346	3	1383	HASVLFCRVMAASKTQGAVARMQEDRDGSC
'		•		-		
						STVGGVGYGDSKDCILEPLSLPESPGGTTTLE
	1]			GSPSVPCIFCEEHFPVAEQDKLLKHMIIEHKIV
!						IADVKLVADFQRYILYWRKRFTEQPITDFCSV
I						IRINSTAPFEEQENYFLLCDVLPEDRILREELQ
İ	l j	- 1	- 1	}		KQRLREILEQQQQERNDTNFHGVCMFCNEEF
		-	ŀ	1	[LGNRSVILNHMAREHAFNIGLPDNIVNCNEFL
1		1	l	1		CTLQKKLDNLQCLYCEKTFRDKNTLKDHMR
			ł	Į		KKQHRKINPKNREYDRFYVINYLELGKSWEE
			1	ļ		VQLEDDRELLDHQEDDWSDWEEHPASAVCL
			J	i		FCEKOAETIEKLYVHMEDAHEFDLLKIKSELG
			ł			LNFYQQVKLVNFIRRQVHQCRCYGCHVKFKS
		ŀ	ļ			KADLRTHMEETKHTSLLPDRKTWDQLEYYFP
	1		ı	ł	ŀ	TYENDTLLWTLSDSESDLTAQEQNENVPIISE
	į			l		
702	2052	A	5356	2500	1540	DTSKLYALKQSSILNQLLL
.02	2002		2220	2502	1540	MAAATRGCRPWGSLLGLLGLVSAAAAAWD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isolcucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LASLRCTLGAFCECDFRPDLPGLECDLAQHL AGQHLAKALVVKALKAFVRDPAPTKPLVLSL HGWTGTGKSYVSSLLAHYLFQGGLRSPRVH HFSPVLHFPHPSHIERYKKDLKSWVQGNLTA
						CGRSLFLFDEMDKMPPGLMEVLRPFLGSSWV VYGTNYRKAIFIFISNTGGEQINQVALEAWRS RRDREEILLQELEPVISRAVLDNPHHGFSNSGI MEERLLDAVVPFLPLQRIIIVRHCVLNELAQL GLEPRDEVVQAVLDSTTFFPEDEQLFSSNGCK TVASRIAFFL
703	2053	A	5380	278	657	LFLQKLRMKTEEEARTHTEIEMFLRKEQQKL EERLEFWMEKYDKDTEMKQNELNALKATKA SDLAHLQDLAKMIREYEQVIIEDRIEKERSKK KVKQDLLELKSVIKLQAWWRGTMIRREIGGF KM
704	2054	A	5381	1	1003	FRGRAVKMAAVVEVEVGGGAAGERELDEV DMSDLSPEEQWRVEHARMHAKHRGHEAMH AEMVLILIATLVVAQLLLVQWKQRHPRSYN MVTLFQMWVVPLYFTVKLHWWRFLVIWILF SAVTAFVTFRATRKPLVQTTPRLVYKWFLLIY KISYATGIVGYMAVMFTLFGLNLLFKIKPEDA MDFGISLLFYGLYYGVLERDFAEMCADYMA STIGFYSESGMPTKHLSDSVCAVCGQQIFVDV SEEGIIENTYRLSCNHVFHEFCIRGWCIVGKK QTCPYCKEKVDLKRMFSNPWERPHVMYGQL LDWLRYLVAWQPVIIGVVQGINYILGLE
705	2055	Α	5396	3	675	IYDRDPLQLATRAGQPLDINMAGEPKPYRPKP GNKRPLSALYRLESKEPFLSVGGYVFDYDYY RDDFYNRLFDYHGRVPPPPRAVIPLKRPRVA VTTTRRGKGVFSMKGGSRSTASGSTGSKLKS DELQTIKKELTQIKTKIDSVLGRLDKIEKQQK AEAEAQKKLLEESLVLIQEECVSEIADHSTEEP AEGGPDADGEEMTDGIEEAFDEDGGHELFLQ IK
706	2056	A	5410	2	98	GRVGLNLEGRGCSEPKWRHCTPTWATEQDSI
707	2057	A	5415	6	287	S PFKLTPSFLSHAFSSGQERKVFIELNHIKKCNT VRGVFVLEEFGNYTILLLGLDSHGSNSNLGAP EEGLGAGRKRTSVEKSGGAGVTRKKRDP
708	2058	Ā	5423	3	291	SSSNPLGSPSTLWKLCSFVLHNKSCCCSFFGS TPTLRAITLTVRVCGFIPEVSKTTNPLGRTNNS GCTIFKTVTLTARSTASLLKSVRPRTHOKE
709	2059	A	5424	679	347	RIRHEEKRGSRGRGRRTSEEDTPKKKKHKGG SEFIDTILSVHPSDVLDMPVDPNEPTYCLCHQ VSYGEMIGCDNPDCPIEWFHFACVDLTTKPK GKWFCPRCVQEKRKKK
710	2060	A	5442	1073	559	QESLKKKIOPKLSLTLSSSVSRGNVSTPPRHSS GSLTPPVTPPITPSSSFRSSTPTGSEYDEEEVDY EESDSDESWTTESAISSEAILSSMCMNGGEEK PFACPVPGCKKRYKNVNGIKYHAKNGHRTQI RVRKPFKCRCGKSYKTAQGLRHHTINFHPPV SAEIIRKMQQ
711	2061	A	5449	1	319	GDSLCVPQYNKYREERVILFLKMASGHAFQP DLVKRIRDAIRMGLSARHVPSLILETKGIPYTL NGKKVEVAVKQIIAGKAVEQGGAFSNPETLD LYRDIPELQGF
712	2062	A	5499	91	749	RPTPGHGDFWMQPLTKDAGMSLSSVTLASAL QVRGEALSEEEIWSLLFLAAEQLLEDLRNDSS DYVVCPWSALLSAAGSLSFQGRVSHIEAAPF

SEQ ID	SEQ ID	Met	SEQ	Desdisted	I D . 2	
NO: of	NO: of	hod	ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in NO.	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-	l	USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	uence	1	09/496	correspondi	to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence	delle	!	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
dence	İ	1	1 214	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
1		1	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	ľ	1	peptide	sequence	/=possible nucleotide deletion, \=possible
				sequence	i	nucleotide insertion
		 -		sequence		MADELL COORDED A GOVERNOV CONTRACTOR
				ļ	ļ	KAPELLQGQSEDEQPDASQMHVYSLGMTLY
	i					WSAGFHVPPHQPLQLCEPLHSILLTMCEDQPH
					1	RRCTLQSVLEACRVHEKEVSVYPAPAGLHIR
713	2063	Ā	5506	22	478	RLVGLVLGTISEVSREPCFSSSSCWSCVAIKI
1,13	2005	^	3300		470	VEELILVSRLDPHLHTPMYFFLAHLSFLDLSFT
		1				TSSIPQLLYNLNGCDKTISYMGCAIQLFLFLGL
						GGVECLLLAVMAYDRCVAICKPLHYMVIMN
						PRLCRGLVSVTWGCGVANSLAMSPVTLRLPR
714	2064	A	5514	25	220	CGHHEVDHFLCEMPALIRMACISTV
' ' 4	2004	^	3314	د ا	220	AIRPYWCENNIIGIGKLSTADGKAFADPEVLR
ł		ł	1 .		}	RLTSSVSCALDEAAAALTRMRAESTANAGQS
715	2065	A	5526	3	010	DK
113	2000	^	2220	3	810	KVTAPRRPQRYSSGHGSDNSSVLSGELPPAM
		İ				GRTALFHHSGGSSGYESLRRDSEATGSASSAP
ł		ŀ	1		ļ	DSMSESGAASPGARTRSLKSPKKRATGLQRR
ł						RLIPAPLPDTTALGRKPSLPGQWVDLPPPLAG
						SLKEPFEIKVYEIDDVERLQRPRPTPREAPTQG
		ļ]			LACVSTRLRLAERRQQRLREVQAKHKHLCEE
						LAETQGRLMLEPGRWLEQFEVDPELEPESAE
						YLAALERATAALEQCVNLCKAHVMMVTCFD
716	2066		7.500			ISVAASAAIPGPQEVDV
716	2066	Α	5529	458	790	SPGYGENKFTVTSXNIAVPLCEMNKIYSYYSD
			•			SSSSERTMDLVLEMCNTNSIHWCGISGRQLG
						KLHPSSSLCLALTLLSSVQGLQSISGLRLTDTF
717	2067		6631		160	LKRTYEYDDIAQVCV
717	2067	A	5531	3	460	NSEDLLKYFNPESWQEDLDNMYLDTPRYRG
				ì		RSYHDRKSKVDLDRLNDDAKRYSCTPRNYS
						VNIREELKLANVVFFPRCLLVQRCGGNCGCG
						TVNWRSCTCNSGKTVKKYHEVLQFEPGHIKR
	2068		7706			RGRAKTMALVDIQLDHHERCDCICSSRPPR
718	2008	Α	5586	311	88	AVLKNMAPMTALGLLDLHILNLILFLSAGEDF
	İ					TSVVSEIMMYILLVFLTLWLLIEMIYCYRKVS
	-00/0					KAEEAAQENA
719	2069	Α]	5598	1	330	KNCANEAVVQKILDRVLSRYDVRLRPNFGSM
i i	ĺ				ľ	LATNSTRGLNEDELMAHGQEKDSSSESEDSC
		1	ľ	i		PPSPGCSFTEGFSFDLLNPDYVPKVDKWSRFL
720	2070		6600		·	FPLAFGLFNIVAAERC
720	2070	Α	5628	798	148	LPPAQIPEAWLLLANVVVVLILVPLKDRLIDP
[]	ľ	ľ	- 1	1	ł	LLLRCKLLPSALQKMALGMFFGFTSVIVAGV
				!		LEMERLHYIHHNETVSQQIGEVLYNAAPLSIW
		.		ļ	1	WQIPQYLLIGISEIFASIPGLEFAYSEAPRSMQG
. [- 1	- 1	1		AIMGIFFCLSGVGSLLGSSLVALLSLPGGWLH
	ľ	ł	ļ	ł		CPKDFGNINNCRMDLYFFLLAGIQAVTALLF
721	2071		F(30	146		VWIAGRYERASQGPASHSRFSRDRG
721	2071	Α	5632	146	536	MSALIVRKLRSAELTLFSELPTVLGANVNAA
	i		1	ŀ	į	KLHETALHHAAKVKNVDLIEMLIEFGGNIYA
	1			l		RDNRGKKPSDYTWSSSAPAKCFEYYEKTPLT
		-		l		LSQLCRVNLRKATGVRGLEKIAKLNIPPRLID
700		<u> </u>				YLSYN
722	2072	A	5638	3	3806	CPSLDIRSEVAELRQLENCSVVEGHLQILLMF
		ł	İ	ĺ	1	TATGEDFRGLSFPRLTQVTDYLLLFRVYGLES
	1	l		l	1	LRDLFPNLAVIRGTRLFLGYALVIFEMPHLRD
ł	ł	ł		l		VALPALGAVLRGAVRVEKNQELCHLSTIDW
	ľ	l				GLLQPAPGANHIVGNKLGEECADVCPGVLGA
	į	l]]		AGEPCAKTTFSGHTDYRCWTSSHCQRVCPCP
	ļ	.		j		HGMACTARGECCHTECLGGCSQPEDPRACV
	1	l	1			ACRHLYFQGACLWACPPGTYQYESWRCVTA
	j	J	j	1		ERCASLHSVPGRASTFGIHQGSCLAQCPSGFT
						RNSSSIFCHKCEGLCPKECKVGTKTIDSIQAA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	}	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	uchic		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
uence		i	914	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of		
1					sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ĺ	1	ĺ	İ	peptide	1	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
	ľ				ŀ	QDLVGCTHVEGSLILNLRQGYNLEPQLQHSL
1			l :			GLVETITGFLKIKHSFALVSLGFFKNLKLIRGD
				ł		AMVDGNYTLYVLDNQNLQQLGSWVAAGLTI
						PVGKIYFAFNPRLCLEHIYRLEEVTGTRGRQN
					İ	KAEINPRTNGDRAACQTRTLRFVSNVTEADRI
1	ŀ				ł	LLRWERYEPLEARDLLSFIVYYKESPFQNATE
j		1		ļ		HVGPDACGTQSWNLLDVELPLSRTQEPGVTL
1	l		}	l		ASLKPWTQYAVFVRAITLTTEEDSPHQGAQS:
		ļ				PIVYLRTLPAAPTVPQDVISTSNSSSHLLVRW
	i i					KPPTQRNGNLTYYLVLWQRLAEDGDLYLND
1						YCHRGLRLPTSNNDPRFDGEDGDPEAEMESD
1	!					CCPCQHPPPGQVLPPLEAQEASFQKKFENFLH
ł	1				·	NAITIPISPWKVTSINKSPORDSGRHRRAAGPL
i	,					RLGGNSSDFEIQEDKVPRERAVLSGLRHFTEY
						RIDIHACNHAAHTVGCSAATFVFARTMPHRE
						ADGIPGKVAWEASSKNSVLLRWLEPPDPNGL
ł						ILKYEIKYRRLGEEATVLCVSRLRYAKFGGV
ł						
ł						HLALLPPGNYSARVRATSLAGNGSWTDSVAF
1			ĺ	,		YILGPEEEDAGGLHVLLTATPVGLTLLIVLAA
						LGFFYGKKRNRTLYASVNPEYFSASDMYVPD
						EWEVPREQISIRELGQGSFGMVYEGLARGLE
ŀ						AGEESTPVALKTVNELASPRECIEFLKEASVM
j						KAFKCHHVVRLLGVVSQGQPTLVIMELMTR
ļ						GDLKSHLRSLRPEAENNPGLPQPALGEMIQM
1						AGEIADGMAYLAANKFVHRDLAARNCMVSQ
						DFTVKIGDFGMTRDVYETDYYRKGGKGLLP
						VRWMAPESLKDGIFTTHSDVWSFGVVLWEIV
				•		TLAEQPYQGLSNEQVLKFVMDGGVLEELEGC
				•		PLQLQELMSRCWQPNPRLRPSFTHILDSIQEEL
						RPSFRLLSFYYSPECRGARGSLPTTDAEPDSSP
						TPRDCSPQNGGPGH
723	2073	Α	5672	1	216	LAWIDNILPEKEKKETDKKRKRKKGAHEDCD
						EEPQFPPPSVIKIPMESVQSDPQNGIHCIARKR
			J			SSSWSYSL
724	2074	A	5704	4235	940	ARGRRSRPVWAASWGGRGRPAARRRPRGLA
						ATMGFELDRFDGDVDPDLKCALCHKVLEDP
						LTTPCGHVFCAGCVLPWVVQEGSCPARCRGR
	·				1	LSAKELNHVLPLKRLILKLDIKCAYATRGCGR
					1	VVKLQQLPEHLERCDFAPARCRHAGCGQVLL
						RRDVEAHMRDACDARPVGRCQEGCGLPLTH
					-	GEQRAGGHCCARALRAHNGALQARLGALHK
						ALKKEALRAGKREKSLVAOLAAAOLELOMT
]						ALRYOKKFTEYSARLDSLSRCVAAPPGGKGE
<u> </u>	•	1	l	ĺ		ETKSLTLVLIRDSGSLGFNIIGGRPSVDNHDG
						SSSEGIFVSKIVDSGPAAKEGGLOIHDRIIEVN
		1				GRDLSRATHDOAVEAFKTAKEPIVVOVLRRT
				l		
]						PRTKMFTPPSESQLVDTGTQTDITFEHIMALT
1 I						KMSSPSPPVLDPYLLPEEHPSAHEYYDPNDYI
i I	' l			1		GDIHQEMDREELELEEVDLYRMNSQDKLGLT
į				ļ		VCYRTDDEDDIGIYISEIDPNSIAAKDGRIREG
)						DRIIQINGIEVQNREEAVALLTSEENKNFSLLI
[ſ	ſ	1	ſ	i	ARAELQLDEGWMDDDRNDFLDDLHMDMLE
				1		EQHHQAMQFTASVLQQKKHDEDGGTTDTAT
1		ļ	1	ļ		ILSNQHEKDSGVGRTDESTRNDESSEQENNG
[[1	1				DDATASSNPLAGQRKLTCSQDTLGSGDLPFS
			1			NESFISADCTDADYLGIPVDECERFRELLELK
				ļ		CQVKSATPYGLYYPSGPLDAGKSDPESVDKE
				ļ		LELLNEELRSIELECLSIVRAHKMQQLKEQYR
Į						ESWMLHNSGFRNYNTSIDVRRHELSDITELPE
[[KSDKDSSSAYNTGESCRSTPLTLEISPDNSLRR

SEQ ID NO: of nucl- eotide	SEQ ID NO: of peptide seq-	Met hod	SEQ ID NO: in USSN	Predicted beginning nucleotide location	Predicted end nucleotide location corresponding	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq- uence	uence		09/496 914	correspondi ng to first	to last amino acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
uchec			3,4	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			İ	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion AAEGISCPSSEGAVGTTEAYGPASKNLLSITE
						DPEVGTPTYSPSLKELDPNQPLESKERRASDG
						SRSPTPSQKLGSAYLPSYHHSPYKHAHIPAHA
						QHYQSYMQLIQQKSAVEYAQSQMSLVSMCK
						DLSSPTPSEPRMEWKVKIRSDGTRYITKRPVR DRLLRERALKIREERSGMTTDDDAVSEMKM
						GRYWSKEERKQHLVKAKEQRRRREFMMOSR
						LDCLKEQQAADDRKEMNILELSHKKMMKKR
						NKKIFDNWMTIQELLTHGTKSPDGTRVYNSF
725	2075	A	5707	3	1770	LSVTTV OISTEVSEAPVANDKPKTLVVKVOKKAADLP
	20.0	••	5.0.			DRDTWKGRFDFLMSCVGYAIGLGNVWRFPY
						LCGKNGGGAFLIPYFLTLIFAGVPLFLLECSLG
[QYTSIGGLGVWKLAPMFKGVGLAAAVLSFW
						LNIYYIVIISWAIYYLYNSFTTTLPWKQCDNP WNTDRCFSNYSMVNTTNMTSAVVEFWERN
						MHQMTDGLDKPGQIRWPLAITLAIAWILVYF
						CIWKGVGWTGKVVYFSATYPYIMLIILFFRGV
						TLPGAKEGILFYITPNFRKLSDSEVWLDAATQ IFFSYGLGLGSLIALGSYNSFHNNVYRDSIIVC
1						CINSCTSMFAGFVIFSIVGFMAHVTKRSIADV
						AASGPGLAFLAYPEAVTQLPISPLWAILFFSM
1. 1						LLMLGIDSQFCTVEGFITALVDEYPRLLRNRR
						ELFIAAVCIISYLIGLSNITQGGIYVFKLFDYYS ASGMSLLFLVFFECVSISWFYGVNRFYDNIQE
1						MVGSRPCIWWKLCWSFFTPIIVAGVFIFSAVQ
				, i		MTPLTMGNYVFPKWGQGVGWLMALSSMVL
						IPGYMAYMFLTLKGSLKQRIQVMVQPSEDIV RPENGPEQPQAGSSTSKEAYI
726	2076	Α	5711	156	423	PRRDPGRTPELRGSAPRKTGANMPVRRGHVA
						PQNTFLGTIIRKFEGQNKKFIIANARVQNCAII
727	2077	Α	5716	3	274	YCNDGFCEMTGFSRPDVMQKPCTCD
121	2077	A	3/10	3	2/4	HASEYFFKLCSFQVFLSFPLATIVIDVGLVVIP LVKSPNVHYVYVLLLVLSGLLFYIPLIHFKIRL
						AWFEKMTCYLQLLFNICLPDVSEE
728	2078	Α	5737	1899	649	IQASRASPYPRVKVDFALSCHEDLLAPISEPIE
1	. 1					WKYHSPEEEISLGPACWLWDFLRRSQQAGFL LPLSGGVDSAATACLIYSMCCQVCEAVRSGN
						EEVLADVRTIVNQISYTPQDPRDLCGRILTTC
					·	YMASKNSSQETCTRARELAQQIGSHHISLNID
						PAVKAVMGIFSLVTGKSPLFAAHGGSSRENL
]						ALQNVQARIRMVLAYLFAQLSLWSRGVHGG LLVLGSANVDESLLGYLTKYDCSSADINPIGG
]				-		ISKTDLRAFVQFCIQRFQLPALQSILLAPATAE
			* 1			LEPLADGQVSQTDEEDMGMTYAELSVYGKL
	ļ				ł	RKVAKMGPYSMFCKLLGMWRHICTPRQVAD KVKRFFSKYSMNRHKMTTLTPAYHAENYSPE
						DNRFDLRPFLYNTSWPWOFRCIENOVLOLER
						AEPQSLDGVD
729	2079	Α	5741	1	5976	PGCAARLSRARAPGPGAAGAGRKRLADPGPP
	ļ					PASRRLRAPGSRPRLAPCTRRAAQPAHARMA PRAAGGAPLSARAAAASPPPFOTPPRCPVPLL
						LLLLLGAARAGALEIQRRFPSPTPTNNFALDG
	(İ	ĺ	AAGTVYLAAVNRLYQLSGANLSLEAEAAVG
.					-	PVPDSPLCHAPQLPQASCEHPRRLTDNYNKIL
	1				l	QLDPGQGLVVVCGSIYQGFCQLRRRGNISAV AVRFPPAAPPAEPVTVFPSMLNVAANHPNAS
					l	TVGLVLPPAAGAGGSRLLVGATYTGYGSSFF
						PRNRSLEDHRFENTPEIAIRSLDTRGDLAKLFT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	J	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		i	l	amino acid residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
					sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
		l		peptide		nucleotide insertion
		<u> </u>		sequence	<u> </u>	FDLNPSDDNILKIKQGAKEQHKLGFVSAFLHP
		[ſ	SDPPPGAQSYAYLALNSEARAGDKESQARSL
			1			LARICLPHGAGGDAKKLTESYIQLGLQCAGG
			}			AGRGDLYSRLVSVFPARERLFAVFERPOGSPA
						ARAAPAALCAFRFADVRAAIRAARTACFVEP
						APDVVAVLDSVVQGTGPACERKLNIQI.QPEQ
			·			LDCGAAHLQHPLSILQPLKATPVFRAPGLTSV
						AVASVNNYTAVFLGTVNGRLLKINLNESMQ
						VVSRRVVTVAYGEPVHHVMQFDPADSGYLY
						LMTSHQMARVKVAACNVHSTCGDCVGAAD
						AYCGWCALETRCTLQQDCTNSSQQHFWTSA
						SEGPSRCPAMTVLPSEIDVRQEYPGMILQISGS
						LPSLSGMEMACDYGNNIRTVARVPGPAFGHQ
						IAYCNLLPRDQFPPFPPNQDHVTVEMSVRVN
			ľ			GRNIVKANFTIYDCSRTAQVYPHTACTSCLSA
ĺ						QWPCFWCSQQHSCVSNQSRCEASPNPTSPQD
						CPRTLLSPLAPVPTGGSQNILVPLANTAFFQG
						AALECSFGLEEIFEAVWVNESVVRCDQVVLH
						TTRKSQVFPLSLQLKGRPARFLDSPEPMTVM
						VYNCAMGSPDCSQCLGREDLGHLCMWSDGC
						RLRGPLQPMAGTCPAPEIRAIEPLSGPLDGGT
						LLTIRGRNLGRRLSDVAHGVWIGGVACEPLP
						DRYTVSEEIVCVTGPAPGPLSGVVTVNASKE
						GKSRDRFSYVLPLVHSLEPTMGPKAGGTRITI
,						HGNDLHVGSELQVLVNDTDPCTELMRTDTSI
						ACTMPEGALPAPVPVCVRFERRGCVHGNLTF
				•		WYMQNPVITAISPRRSPVSGGRTITVAGERFH
						MVQNVSMAVHHIGREPTLCKVLNSTLITCPSP
				•		GALSNASAPVDFFINGRAYADEVAVAEELLD
					ŀ	PEEAQRGSRFRLDYLPNPQFSTAKREKWIKH
	J					HPGEPLTLVIHVSTKGAGKEQDSLGLQSHEY
						RVKIGQVSCDIQIVSDRIIHCSVNESLGAAVGQ
						LPITIQVGNFNQTIATLQLGGSETAIIVSIVICSV
						LLLLSVVALFVFCTKSRRAERYWQKTLLQME
]	J	į	EMESQIREEIRKGFAELQTDMTDLTKELNRSQ GIPFLEYKHFVTRTFFPKCSSLYEERYVLPSOT
					-	LNSQGSSQAQETHPLLGEWKIPESCRPNMEE
				ļ		GISLFSSLLDNKHFLIVFVHALEQQKDFAVRD
				•	ļ	RCSLASLLTIALHGKLEYYTSIMKELLVDLID
ł	}				ļ	ASAAKNPKLMLRRTESVVEKMLTNWMSICM
						YSCLRETVGEPFFLLLCAIKQQINKGSIDAITG
				1		KARYTLNEEWLLRENIEAKPRNLNVSFOGCG
İ			[[1		MDSLSVRAMDTDTLTQVKEKILEAFCKNVPY
] [ļ		SQWPRAEDVDLEWFASSTQSYILRDLDDTSV
				- 1	-	VEDGRKKLNTLAHYKIPEGASLAMSLIDKKD
J	ļ			ŀ		NTLGRVKDLDTEKYFHLVLPTDELAEPKKSH
				į		ROSHRKKVLPEIYLTRLLSTKGTLOKFLDDLF
į	1				ļ	KAILSIREDKPPLAVKYFFDFLEEQAEKRGISD
] {	İ		PDTLHIWKTNSLPLRFWVNILKNPQFVFDIDK
1	1		}	1		TDHIDACLSVIAQAFIDACSISDLQLGKDSPTN
l	Ì					KLLYAKEIPEYRKIVQRYYKQIQDMTPLSEOE
- 1	1					MNAHLAEESRKYQNEFNTNVAMAEIYKYAK
l				į		RYRPQIMAALEANPTARRTQLQHKFEQVVAL
						MEDNIYECYSEA
730	2080	A	5744	3	292	QPSPLFHSHLETLQLLRTAQLPEQVSWPWGQ
-				ļ]	VANGKGNQRNMGSPQPSLLAFERNLELQIMG
]			.	1		LGYSLLMGKLRPRVAKDTLRVHRDSTPSPLT
						LKD
731	2081	A	5747	1	382	FLKCMRKAFRSSKLLQVGYTPDGKDDYRWC
						FRVDEVNWTTWNTNVGIINEDPGNCEGVKRT

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first	to last amino acid residue	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
				amino acid residue of peptide	of peptide sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion LSFSLRSSRVSGRHWKNFALVPLLREASARD
732	2082	Α	5753	198	3	RQSAQPEEVYLRQFSGSLKPEDAEVFKSPAAS GEK AQAESSTVASPEATAGPLCTRIPNVPPPTPIRP
733	2083	Λ	5754	2	2223	PGKLQAQLPCPSPVRFTSARIPPASRPQTKS AAGPPGLEAEGRAPESAGPGPGGDAAETPGL
				_		PPAHSGTLMMAFRDVTVQIANQNISVSSSTAL SVANCLGAQTVQAPAEPAAGKAEQGETSGR EAPEAPAVGREDASAEDSCAEAGASGAADG ATAPKTEEEEEEEEETAEVGRGAEAEAGDLEQ LNRTSTSTKSAKSGSEASASASKDALQAMILS
						LPRYHCENPASCKSPTLSTDTLRKRLYRIGLN LFNINPDKGIQFLISRGFIPDTPIGVAHFLLQRK GLSRQMIGEFLGNSKKQFNRDVLDCVVDEM DFSSMELDEALRKFQAHIRVQGEAQKVERLIE
						AFSQRYCMCNPEVVQQFHNPDTIFILAFAIILL NTDMYSPNIKPDRKMMLEDFIRNLRGVDDG ADIPRELVVGIYERIQQKELKSNEDHVTYVTK VEKSIVGMKTVLSVPHRRLVCCSRLFEVTDV
						NKLQKQAAHQREVFLFNDLLVILKLCPKKKS SSTYTFCKSVGLLGMQFQLFENEYYSHGITLV TPLSGSEKKQVLHFCALGSDEMQKFVEDLKE
				•		SIAEVTELEQIRIEWELEKQQGTKTLSFKPCGA QGDPQSKQGSPTAKREAALRERPAESTVEVSI HNRLQTSQHNSGLGAERGAPVPPPDLQPSPPR QQTPPLPPPPPTPPGTLVQCQQIVKVIVLDKPC
	:					LARMEPLLSQALSCYTSSSSDSCGSTPLGGPG SPVKVTHQPPLPPPPPPYNHPHQFCPPGSLLH GHRYSSGSRSLV
734	2084	A	5788	8	362	SSVMGDLVGQGLEEQIVARDENSWLIDGGTP IDDVMRVLDIDEFPQSGNYETIGGFMMFMLR KIPKRTDSVKFAGYKFEVVDIDNYRIDQLLVT RIDSKATALSPKLPDAKDKEESVA
735	2085	A	5827	1	1257	MVFSAVLTAFHTGTSNTTFVVYENTYMNITL PPPFQHPDLSPLLRYSFETMAPTGLSSLTVNST AVPTTPAAFKSLNLPLQITLSAIMIFILFVSFLG
						NLVVCLMVYQKAAMRSAINILLASLAFADM LLAVLNMPFALVTILTTRWIFGKFFCRVSAMF FWLFVIEGVAILLIISIDRFLIIVQRQDKLNPYR
				·		AKVLIAVSWATSFCVAFPLAVGNPDLQIPSRA PQCVFGYTTNPGYQAYVILISLISFFIPFLVILY SFMGILNTLRHNALRIHSYPEGICLSQASKLGL MGLQRPFQMSIDMGFKTRAFTTILILFAVFIVC
		-				WAPFTTYSLVATFSKHFYYQHNFFEISTWLL WLCYLKSALNPLIYYWRIKKFHDACLDMMP KSFKFLPQLPGHTKRRIRPSAVYVCGEHRTVV
736	2086	A	5870	3	268	FTRSDELARHYRTHTGEKRFSCPLCPKQFSRS DHLTKHARRHPTYHPDMIEYRGRRRTPRIDPP LTSEVESSASGSGPGPAPSFTTCL
737	2087	A	5871	2	521	LTWPQLFLETLPELLHMSRPAEDGPSPGALVR RSSSLGYISKAEEYFLLKSRSDLMFEKQSERH GLARRLTTARRPPASSEQAQQELFNELKPAV DGANFIVNHMRDQNNYNEEKDSWNRVART VDRLCLFVVTPVMVVGTAWIFLQGVYNQPPP
738	2088	A	5881	1	1160	QPFPGDPYSYNVQDKRFI LVVTAITAILAFPNEYTRMSTSELISELFNDCG LLDSSKLCDYENRFNTSKGGELPDRPAGVGV YSAMWQLALTLILKIVITIFTFGMKIPSGLFIPS MAVGAIAGRLLGVGMEQLAYYHQEWTVFNS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	Ì	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ļ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		İ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			}	peptide		/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
				ļ		WCSQGADCITPGLYAMVGAAACLGGVTRMT
						VSLVVIMFELTGGLEYIVPLMAAAMTSKWVA
1				İ		DALGREGIYDAHIRLNGYPFLEAKEEFAHKTL
		ĺ				AMDVMKPRRNDPLLTVLTQDSMTVEDVETII SETTYSGFPVVVSRESQRLVGFVLRRDLIISIE
1						NARKKQDGVVSTSIIYFTEHSPPLPPYTPPTLK
}		İ				LRNILDLSPFTVTDLTPMEIVVDIFRKLGLROC
					,	LVTHNGRLLGIITKKDVLKHIAQMANQDPDSI
						LFN
739	2089	Α	5892	2	916	TLQLAASVPFFAISLISWWLPESARWLIINGKP
		' '		-		DQALQELRKVARINGHKEAKNLTIEVLMSSV
						KEEVASAKEPRSVLDLFCVPVLRWRSCAMLV
						VNFSLLISYYGLVFDLQSLGRDIFLLQALFGA
						VDFLGRATTALLLSFLGRRTIQAGSQAMAGL
				į		AILANMLVPQDLQTLRVVFAVLGKGCFGISL
						TCLTIYKAELFPTPVRMTADGILHTVGRLGA
						MMGPLILMSRQALPLLPPLLYGVISIASSLVVL
						FFLPETQGLPLPDTIQDLESQKSTAAQGNRQE
L						AFTVESTSLLEIVALHGAL
740	2090	Α	5900	2	426	RPIKTLGIGFHFSVDGVHFLTQREVQNLWKE
1						NLIILDTAKKHGYEVVDTFTITMGRYKEFLQG
						KCGCHFHEVVKSKLSKEYNFIKMKRSRNHIM
						GRYFSNQSKLQQGTVTNFRSPYHVRGPINQV
741	2001	_	5010		410	CSEILLSRMCANKRTM
741	2091	A	5910	3	412	RMPESTLLIICENGYILEAPLPTIKQEEDDHDV
						VSYEIKDMCIKCFHFSSVKSKILRLIEIEKRER
						QRELKEKIREERRNKLAAEMGEDGEKEFQEE EEEKEEEEEEEPLPEIFIPSTPSPILCGFYSEPG
					:	KFWV
742	2092	A	5936	1	482	MGCRLLCCVVFCLLQAGPLDTAVSQTPKYLV
' '	20,2	·	3,50	•	402	TQMGNDKSIKCEQNLGHDTMYWYKQDSKK
						FLKIMFSYNNKELIINETVPNRFSPKSPDKAHL
						NLHINSLELGDSAVYFCASSQDTALQSHCIPV
1		1				HKPPGSARKLQGSVCTCTQGSSLHSLMASDG
[[i	1	ĺ			VPVC
743	2093	Α	5938	1	1566	MNSFFGTPAASWCLLESDVSSAPDKEAGRER
						RALSVQQRGGPAWSGSLEWSRQSAGDRRRL
						GLSRQTAKSSWSRSRDRTCCCRRAWWILVPA
}		J				ADRARRERFIMNEKWDTNSSENWHPIWNVN
		Ì		ł		DTKHHLYSDINITYVNYYLHQPQVAAIFIISYF
	j	1				LIFFLCMMGNTVVCFIVMRNKHMHTVTNLFI
		ļ	1	j	j	LNLAISDLLVGIFCMPITLLDNIIAGWPFGNTM
	1	į	[1		CKISGLVQGISVAASVFTLVAIAVDRFQCVVY
	}		ļ			PFKPKLTIKTAFVIIMIIWVLAITIMSPSAVMLH
			ľ			VQEEKYYRVRLNSQNKTSPVYWCREDWPNQ
		1		j		EMRKIYTTVLFANIYLAPLSLIVIMYGRIGISLF
		ł	ļ	J		RAAVPHTGRKNQEQWHVVSRKKQKIIKMLLI
	ļ	ļ	1			VALLFILSWLPLWTLMMLSDYADLSPNELQII
]	1	l			NIYIYPFAHWLAFGNSSVNPIIYGFFNENFRRG
]			ŀ		FQEAFQLQLCQKRAKPMEAYALKAKSHVLIN
	1	- 1	ĺ	ſ	[TSNQLVQESTFQNPHGETLLYRKSAEKPQQE
744	2094	A	5966	149	327	LVMEELKETTNSSEI SHVCVSHYAGSSGCPAGAGAGAVALGISAVA
'**	2034	^	J300	147	321	
745	2095	A	5970	413	856	LYDYQGGRLGVARGAWYMEAPDIRQGDM GAPHTDWAWAPTPMSGLGSGRGRQGTLASS
""	2073	^	3710	71.5	0.00	PLSLPLLLAGVTGILATELFDQMARPAACMV
. l	[l	1		ļ	CGALMWIMLILVGLGFPFIMEALSHFLYVPFL
i i	ļ	l	l			GVCVCGAIYTGLFLPETKGKTFQEISKELHRL
J	J	J	j		ļ	NFPRRAQGPTWRSLEVIQSTEL
	1				i	

SEO ID	SEQ ID	Met	SEQ	Predicted	Desdisted and	T A - i (A - A) - i C - C
NO: of	NO: of	hod			Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
		nou	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ļ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	ncuce	i	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	ŀ			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	1	İ	peptide		/=possible nucleotide deletion, \=possible
	1	Į.		sequence	į.	nucleotide insertion
746	2096	A	5971	3	1242	
/40	2090	ΙΛ.	39/1	3	1343	AQTARRIIGLELDTEGHRLFVAFSGCIVYLPLS
			1	i		RCARHGACQRSCLASQDPYCGWHSSRGCVDI
}	1	1	1	ŀ]	RGSGGTDVDQAGNQESMEHGDCQDGATGSQ
	1	i			[SGPGDSAYGVRRDLPPASASRSVPIPLLLASV
1			İ			AAAFALGASVSGLLVSCACRRAHRRRGKDIE
ı						TPGLPRPLSLRSLARLHGGGPEPPPPSKDGDA
	İ	l	1 :			VQTPQLYTTFLPPPEGVPPPELACLPTPESTPE
		1			ļ	LPVKHLRAAGDPWEWNQNRNNAKEGPGRSR
ł	l	1			ĺ	COLL + COD + DDIA + TO TO TO COLO + 1 TO TO
						GGHAAGGPAPRVLVRPPPPGCPGQAVEVTTL
		ĺ	1			EELLRYLHGPQPPRKGAEPPAPLTSRALPPEP
	1	l	ļ		ļ	APALLGGPSPRPHECASPLRLDVPPEGRCASA
1	1		1			PARPALSAPAPRLGVGGGRRLPFSGHRAPPAL
1	1	Į				LTRVPSGGPSRYSGGPGKHLLYLGRPEGYRG
	1					RALKRYDVEKPQLSLKPPLVGPSSRQAVPNG
ł	ł	ł	1			GRFNF
747	2097	Ā	5998	2	754	DHASLPCSWNHRFDVETRHVFIGDHSGOVTI
1 ' ''	2037	**	3370	-	754	
	•	1				LKLEQENCTLVTTFRGHTGGVTALCWDPVQ
	1		1			RVLFSGSSDHSVIMWDIGGRKGTAIELQGHN
	ļ]]			DRVQALSYAQHTRQLISCGGDGGIVVWNMD
i						VERQETPEWLDSDSCQKCDQPFFWNFKQMW
			1			DSKKIGLRQHHCRKCGKAVCGKCSSKRSSIPL
		1				MGFEFEVRVCDSCHEAITDEERAPTATFHDSK
ł	1	İ	! !			HNIVHVHFDATRGWLLTSGTDKVIKLWDMT
			i I			PVVS
748	2098	A	6001	2	747	AMVFGGVVPYVPQYRDIRRTONADGFSTYV
' ' '		١.,	0001	-	, -,	
l			1			CLVLLVANILRILFWFGRRFESPLLWQSAIMIL
l	j		J j			TMLLMLKLCTEVRVANELNARRRSFTAADS
						KDEEVKVAPRRSFLDFDPHHFWQWSSFSDYV
			 	•		QCVLAFTGVAGYITYLSIDSALFVETLGFLAV
	į.		! i			LTEAMLGVPQLYRNHRHQSTEGMSIKMVLM
i	1		1	Į.		WTSGDAFKTAYFLLKGAPLQFSVCGLLQVLV
	1		i I			DLAILGQAYAFARHPQKPAPHAVHPTGTKAL
749	2099	A	6002	2	447	GRPDRSELVRMHILEETFAEPSLQATQMKLK
ł				_		RARLADDLNEKIAORPGPMELVEKNILPVDSS
				1		VKEAIIGVGKEDYPHTQGDFSFDEDSSDALSP
}]		1	j		
1				ĺ	[DQPASQESQGSAASPSEPKVSESPSPVTTNTP
750	2100	_	(001		40.5	AQFASVSPTVPEFLKTPPTAD
750	2100	A	6004	2	427	LLTQAMLVLPHRPQWFTPGPRLQAQGPCQEG
}		İ		- 1		WRWELRLRNYVPEDEDLNKRRVPQAKPDAV
			l	- 1	ļ	QEKVKEQLEAAKPEPVIEEVDLAKLAPRKPD
]	ļ	WDLKRDVAKKLEKLLKRTQRAIAELIRERLK
]			1		GQEDSLDSAVDAATEHKTC
751	2101	A	6007	33	1280	TDQAKVDNQPEKLVRSAEDVSTVPTQPDNPF
		••	"""			CHEDRI RDIVERGRADIA EL ODECIDADE CANALITA DE LA CAMBRIA DE LA CAMBRIA DE LA LIGIDA DE LA CAMBRIA DE
						SIPDKLKRMSKSVPAFLQDESDDRETDTASE
			- 	l		SSYQLSRHKKSPSSLTNLSSSSGMTSLSSVSGS
			[l		VMSVYSGDFGNLEVKGNIQFAIEYVESLKEL
			}	I	ļ	HVFVAQCKDLAAADVKKQRSDPYVKAYLLP
				-	I	DKGKMGKKKTLVVKKTLNPVYNEILRYKIEK
			1	i	ľ	QILKTQKLNLSIWHRDTFKRNSFLGEVELDLE
			ŀ	l	ľ	TWDWDNKQNKQLRWYPLKRKTAPVALEAE
	} }		ļ	. 1	1	NRGEMKLALQYVPEPVPGKKLPTTGEVHIWV
				!	[
				į	j	KECLDLPLLRGSHLNSFVKCTILPDTSRKSRQ
				Į.	i	KTRAVGKTTNPIFNHTMVYDGFRPEDLMEAC
				l	ļ	VELTVWDHYKLTNQFLGGLRIGFGTGKSYGT
				l	i	EVDWMDSTSEEVALWEKMVNSPNTWIEATL
						PLRMLLIAKISK
752	2102	Α	6028	108	1283	KEIFSPFELISVKPLCLLLGVTCSQSMAFEELL
		i		1	i	SQVGGLGRFQMLHLVFILPSLMLLIPHILLENF
				Ì	1	AAAIPGHRCWVHMLDNNTGSGNETGILSEDA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LLRISIPLDSNLRPEKCRRFVHPQWQLLHLNG TIHSTSEADTEPCVDGWVYDQSYFPSTIVTKW DLVCDYQSLKSVVQFLLLTGMLVGGIIGGHV
						SDRFGRRFILRWGLLQLAITDTCAAFAPTFPV YCVLRFLAGFSSMIIISNNSLPITEWIRPNSKAL VVILSSGALNIGQIILGGLAYVFRDWQTLHVV ASVPFFVFFLLSRWLVESARWLIITNKLDEGL KALRKVARTNGIKNAEETLNIEVVRSTMQEE LDAAQTKTTVWDLFRNPSMRKRICILVFLRK KNLKEKA
753	2103	A	6043		1470	DSFESILRLIFEIHHSGEKGDIVVFLACEQDIEK VCETVYQGSNLNPDLGELVVVPLYPKEKCSL FKPLDETEKRCQVYQRRVVLTTSSGEFLIWSN SVRFVIDVGVERRKVYNPRIRANSLVMQPISQ SQAEIRKQILGSSSSGKFFCLYTEEFASKDMTP LKPAEMQEANLTSMVLFMKRIDIAGLGHCDF MNRPAPESLMQALEDLDYLAALDNDGNLSE FGIIMSEFPLDPQLSKSILASCEFDCVDEVLTIA AMVTAPNCFSHVPHGAEEAALTCWKTFLHPE GDHFTLISIYKAYQDTTLNSSSEYCVEKWCRD YFLNCSALRMADVIRAELLEIIKRIELPYAEPA FGSKENTLNIKKALLSGYFMGIARDVDGSGN YLMLTHKQVAQLHPLSGYSITKKMPEWVLF HKFSISENNYIRITSEISPELFMQLVPQYYFSNL PPSESKDILQQVVDHLSPVSTMNKEQQMCET CPETEQRCTLQ
754	2104	A	6055	2	394	YYALHHWPFPDLLCQTTGAIFQMNMYGSCIF LMLINVDRYAAIVHPLRLRHLRRPRVARLLC LGVWALILVFAVPAARVHRPSRCRYRDLEVR LCFESFSDELWKGRLLPLVLLAEALGFLLPLA AVVYSS
755	2105	A	6059	3	1795	LGLGSGTLLSVSEYKKKYREHVLQLHARVKE RNARSVKITKRPTKLLIAPESAAPEEALGPAEE PEPGRARRSDTHTFNRLFRRDEEGRRPLTVVL QGPAGIGKTMAAKKILYDWAAGKLYQGQVD FAFFMPCGELLERPGTRSLADLILDQCPDRGA PVPQMLAQPQRLLFILDGADELPALGGPEAAP CTDPFEAASGARVLGGLLSKALLPTALLLVTT RAAAPGRLQGRLCSPQCAEVRGFSDKDKKK YFYKFFRDERRAERAYRFVKENETLFALCFV PFVCWIVCTVLRQQLELGRDLSRTSKTTTSVY LLFITSVLSSAPVADGPRLQGDLRNLCRLARE GVLGRRAQFAEKELEQLELRGSKVQTLFLSK KELPGVLETEVTYQFIDQSFQEFLAALSYLLE DGGVPRTAAGGVGTLLRGDAQPHSHLVLTT RFLFGLLSAERMRDIERHFGCMVSERVKQEA LRWVQGQGGCPGVAPEVTEGAKGLEDTEE PEEEEEGEEPNYPLELLYCLYETQEDAFVRQA LCRFPELALQRVRFCRMDVAVLSYCVRCCPA GQALRLISCRLVAAQEKKKKSLGKRLQASLG GG
756	2106	A .	6060	12	436	SGRPTRPAKPTGQGMGRFMLTLVCQGSIMMS ARDLIMNNLTELQPGLFHHLRFLEELRLSGNH LSHIPGQAFSGLYSLKILMLHNNQLGGIPAQA LWELPSLQSLRLDANLISLVPERSFEGLSSLRH LWLDDNALTEIPS
757	2107	A	6063	54	419	ITPLGLGAADMCAFPWLLLLLLQEGSQRRL WRWCGSEEVVAVLQESISLPLEIPPDEEVENII WSSHKSLATVVPGKEGHPATIMVTNPHYQG

NO: of nucleotide cotide sequence ue	
Corticle sequence USSN U	
sequence uence 09/496 correspondi ng to first amino acid residue of peptide residue of peptide sequence vence	
1914 ng to first amino acid residue of peptide residue of peptide residue of peptide sequence 2=Glutamine, N=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Stop codon, /=possible nucleotide deletion, \=p	
residue of peptide sequence	
peptide sequence /=possible nucleotide deletion, >possible nucleotide insertion QILTMLLRSLQQPSASWPRDCSSSCSW QISCPATIFVPMFSHSLIGIGEEYQLPYY PSDPSYEDMREVVCVKRLPIVSNRWN LRAVLKLMSECWAHNPASRLTALRIKK MVESQDVKI RESPENSE SEQUENCE SKDPSSKSGNLLETSEVGWTSNPEELDPY LLGKSGLSCQVGSATSHPVSCQEPIDEDD KDKSTAGREFSGQVSHQTTSENQCTPIP HSSVADMQNMPAAVHALLTQPSLSAAP RYLGTLPSTGSTTLPQCHAGNATVW FROM FRO	
Sequence nucleotide insertion	
758 2108	
Total	
PSDPSYEDMREVVCVKRLRPIVSNRWN: LRAVLKLMSECWAHNPASRLTALRIKK: MVESQDVKI 759 2109 A 6072 3 650 PGRFRFAALEERAMEKLREKVPFQNRG LSSIIPNNSDTRKATETTSLSSKPEYVNPI SKDPSSKSGNLLETSEVGWTSNPEELDP! LLGKSGLSCQVGSATSIPVSCQEPIDEDG KDKSTAGREFSGQVSHQTTSENQCTPIP: HSSVADMQNMPAAVHALLTQPSLSAAP RYLGTLPSTGSTTLPQCHAGNATVW 760 2110 A 6077 3 730 PLRLTLMEEVLLLGLKDREGYTSFWNDG LRGCMLIELPLRGRLQLEACGMRRKSLL VICKSDAPTGDVLLDEALKHVKETQPPE NWIELLSGETWNPLKLHYQLRNVRERL VEKGVLTTEKQNFLLFDMTTHPLTNNNI LIKKVQEAVLDK WVNDPHRMDRRLLAI AHASDVLENAFAPLLDEQYDLATKRVR LDPEVECLKANTINEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSR YMEL YDLNRDLLNG YKIRWI TELLGNLKA VNQAIQRAGRLRVGKPKNG ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAÏTCGEKEKQDLIKSLAMLKDGFFTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLYEEAKRRIAN	
LRAVLKLMSECWAHNPASRLTALRIKK MVESQDVK1 759 2109 A 6072 3 650 PGRRFRPAALEERAMEKLREKVPFQNRK LSSIIPNNSDTRKATETTSLSSKPEYVNPI SKDPSSKSGNLLETSEVGWTSNPEELDP LLGKSGLSCQVGSATSHPVSCQEPIDEDO KDKSTAGREFSGQVSHQTTSENQCTPIPS HSSVADMQNMPAAVHALLTQPSLSAAP RYLGTLPSTGSTTLPQCHAGNATVW 760 2110 A 6077 3 730 PLRLTLMEEVLLLGLKDREGYTSFWND LRGCMLIELPLRGRLQLEACGMRRKSLL VICKSDAPTGDVLLDEALKHVKETQPPE NWIELLSGETWNPLKLHYQLRNVRERL VEKGVLTTEKQNFLLFDMTTHPLTNNN LIKKVQEAVLDK WVNDPHRMDRRLLAI AHASDVLENAFAPLLDEQYDLATKRVR LDPEVECLKANTNEVLWAVVAAFTK LDPEVECLKANTNEVLWAVVAAFTK HOT HOT HOT HOT A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARL MKTMKSRYMELYDLNRDLLNGYKIRW TELLGNLKAVNQAIQRAGRLRVGKPKN ACRDAIRSNNINTLFKIMRVGTASS ACRDAIRSNNINTLFKIMRVGTASS HSDLWSSSSSLESSSFPLPKQYLDVSSQT SFGINSNNQLAEKVRLRLYEEAKRRIAN SFGINSNNQLAEKVRLRLYEEAKRRIAN SFGINSNNQLAEKVRLRLYEEAKRRIAN STGINSNNNQLAEKVRLRLYEEAKRRIAN STGINSNNNQLAEKVRLRLYEEAKRRIAN LRAVLKEKVPFQNR LLGKSGLSCAUCHTSLAMEKVPT HSDLWSSSSLESSSFPLPKQYLDVSSQT SFGINSNNQLAEKVRLRLYEEAKRRIAN STGINSNNNQLAEKVRLRLYEEAKRRIAN STGINSNNNQLAEKVRLRLYEEAKRRIAN STGINSNNNQLAEKVRLRLYEEAKRRIAN STGINSNNNQLAEKVRLRLYEEAKRRIAN STGINSNNNQLAEKVRLRLYEEAKRRIAN STGINSNNNGLAEKVRLRLYEEAKRRIAN STGINSNN STGINSN STGINSN STGINSNN NNSDTRKATETTSLSSKPEYVNPI SKDPSSKSGNLLETSEVGWTSNPEELDPI LLGKSGLSCQVGSATSHPVSCQEPIDEDG KDKSTAGREFSGQVSHQTTSENQCTPIPS HSSVADMQNMPAAVHALLTQPSLSAAP RYLGTLPSTGSTTILPQCHAGNATVW 760 2110 A 6077 3 730 PLRLTLMEEVLLLGLKDREGYTSFWNDG LRGCMLIELPLRGRLQLEACGMRRKSLI VICKSDAPTGDVLLDEALKHVKETQPPE NWIELLSGETWNPLKLHYQLRNVRERL VEKGVLTTEKQNFLLFDMTTHPLTNNNI LIKKVQEAVLDKWVNDPHRMDRRLLAI AHASDVLENAFAPLLDEQYDLATKRVR LDPEVECLKANTNEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARL MKTMKSRYMELYDLNRDLLNGYKIRW TELLGNLKAVNQAIQRAGRLRVGKPKNG ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLYEEAKRRIAN	DEC
759 2109 A 6072 3 650 PGRRFRPAALEERAMEKLREKVPFQNRG LSSIIPNNSDTRKATETTSLSSKPEYVNPI SKDPSSKSGNLLETSEVGWTSNPEELDPI LLGKSGLSCQVGSATSHPVSCQEPIDEDO KDKSTAGREFSGQVSHQTTSENQCTPIPS HSSVADMQNMPAAVHALLTQPSLSAAP RYLGTLPSTGSTTLPQCHAGNATVW 760 2110 A 6077 3 730 PLRLTLMEEVLLLGLKDREGYTSFWNDO LRGCMLIELPLRGRLQLEACGMRRKSLL VICKSDAPTGDVLLDEALKHVKETQPPE NWIELLSGETWNPLKLHYQLRNVRERLA VEKKVQEAVLDK WVNDPHRMDRRLLAI AHASDVLENAFAPLLDEQYDLATKRVR LDPEVECLKANTNEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMELYDLNRDLLNGYKRWI TELLGNLKAVNQAIQRAGRLRVGKPKNI ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQT SFGINSNNQLAEKVRLRLRYEEAKRRIAN	ILAK
LSSIPNNSDTRKATETTSLSSKPEYVNPI SKDPSSKSGNLLETSEVGWTSNPEELDPI LLGKSGLSCQVGSATSHPVSCQEPIDEDG KDKSTAGREFSGQVSHQTTSENQCTPIPS HSSVADMQNMPAAVHALLTQPSLSAAP RYLGTLPSTGSTTLPQCHAGNATVW 760 2110 A 6077 3 730 PLRLTLMEEVLLLGLKDREGYTSFWNDG LRGCMLIELPLRGRLQLEACGMRRKSLL VICKSDAPTGDVLLDEALKHVKETQPPE NWIELLSGETWNPLKLHYQLRNVRERLA VEKGVLTTEKQNFLLFDMTTHPLTINNNI LIKKVQEAVLDK WVNDPHRMDRRLLAI AHASDVLENAFAPLLDEQYDLATKRVR LDPEVECLKANTNEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMELYDLNRDLLNGYKIRWI TELLGNLKAVNQAIQRAGRLRVGKPKNI ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQT SFGINSNNQLAEKVRLRLRYEEAKRRIAN	W.Carr
SKDPSSKSGNLLETSEVGWTSNPEELDPI LLGKSGLSCQVGSATSHPVSCQEPIDEDG KDKSTAGREFFSGQVSHQTTSENQCTPIPS HSSVADMQNMPAAVHALLTQPSLSAAP RYLGTLPSTGSTTLPQCHAGNATVW 760 2110 A 6077 3 730 PLRLTLMEEVLLIGLKDREGYTSFWNDG LRGCMLIELPLRGRLQLEACGMRRKSLL VICKSDAPTGDVLLDEALKHVKETQPPE NWIELLSGETWNPLKLHYQLRNVRERLA VEKGVLTTEKQNFLLFDMTTHPLTINNNI LIKKVQEAVLDK WVNDPHRMDRRLLAI AHASDVLENAFAPLLDEQYDLATKRVR LDPEVECLKANTNEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMELYDLNRDLLNGYKIRWI TELLGNLKAVNQAIQRAGRLRVGKPKNI ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	
LLGKSGLSCQVGSATSHPVSCQEPIDEDOK KDKSTAGREFSGQVSHQTTSENQCTPIPS HSSVADMQNMPAAVHALLTQPSLSAAP RYLGTLPSTGSTTLPQCHAGNATVW 760 2110 A 6077 3 730 PLRLTLMEEVLLLGLKDREGYTSFWNDG LRGCMLIELPLRGRLQLEACGMRRKSLL VICKSDAPTGDVLLDEALKHVKETQPPE NWIELLSGETWNPLKLHYQLRNVRERLA VEKGVLTTEKQNFLLFDMTTHPLTINNNI LIKKVQEAVLDK WVNDPHRMDRRLLAI AHASDVLENAFAPLLDEQYDLATKRVR LDPEVECLKANTINEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMELYDLNRDLLNGYKIRWI TELLGNLKAVNQAIQRAGRLRVGKPKNI ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	DIA
KDKSTAGREFSGQVSHQTTSENQCTPIPS HSSVADMQNMPAAVHALLTQPSLSAAP RYLGTLPSTGSTTLPQCHAGNATVW 760 2110 A 6077 3 730 PLRLTLMEEVLLLGLKDREGYTSFWNDG LRGCMLIELPLRGRLQLEACGMRRKSLL VICKSDAPTGDVLLDEALKHVKETQPPE NWIELLSGETWNPLKLHYQLRNVRERLA VEKGVLTTEKQNFLLFDMTTHPLTNNNI LIKKVQEAVLDK WVNDPHRMDRRLLAI AHASDVLENAFAPLLDEQYDLATKRVR LDPEVECLKANTNEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMELYDLNRDLLNGYKIRW TELLGNLKAVNQAIQRAGRLRVGKPKNG ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	DRICE
HSSVADMQNMPAAVHALLTQPSLSAAP RYLGTLPSTGSTTLPQCHAGNATVW	STV
760 2110 A 6077 3 730 PLRLTLMEEVLLLGLKDREGYTSFWNDG LRGCMLIELPLRGRLQLEACGMRRKSLL VICKSDAPTGDVLLDEALKHVKETQPPE NWIELLSGETWNPLKLHYQLRNVRERLA VEKGVLTTEKQNFLLFDMTTHPLTNNNI LIKKVQEAVLDK WVNDPHRMDRRLLAI AHASDVLENAFAPLLDEQYDLATKRVR LDPEVECLKANTNEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMELYDLNRDLLNGYKIRWI TELLGNLKAVNQAIQRAGRLRVGKPKNG ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	FAO
LRGCMLIELPLRGRLQLEACGMRRKSLL VICKSDAPTGDVLLDEALKHVKETQPPE NWIELLSGETWNPLKLHYQLRNVRERLA VEKGVLTTEKQNFLLFDMTTHPLTINNNI LIKKVQEAVLDK WVNDPHRMDRRLLAI AHASDVLENAFAPLLDEQYDLATKRVRE LDPEVECLKANTNEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMELYDLNRDLLNGYKIRWI TELLGNLKAVNQAIQRAGRLRVGKPKNI ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	-
VICKSDAPTGDVLLDEALKHVKETQPPE NWIELLSGETWNPLKLHYQLRNVRERLA VEKGVLTTEKQNFLLFDMTTHPLTINNNI LIKKVQEAVLDK WVNDPHRMDRRLLAI AHASDVLENAFAPLLDEQYDLATKRVR LDPEVECLKANTNEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMEL YDLNRDLLNGYKIRWI TELLGNLKAVNQAIQRAGRLRVGKPKNI ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	
NWIELLSGETWNPLKLHYQLRNVRERLA VEKGVLTTEKQNFLLFDMTTHPLTINNNI LIKKVQEAVLDK WVNDPHRMDRRLLAI AHASDVLENAFAPLLDEQYDLATKRVRI LDPEVECLKANTNEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMELYDLNRDLLNGYKIRWI TELLGNLKAVNQAIQRAGRLRVGKPKNI ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	
VEKGVLTTEKQNFLLFDMTTHPLTNNNI LIKKVQEAVLDK WVNDPHRMDRRLLAI AHASDVLENAFAPLLDEQYDLATKRVRI LDPEVECLKANTNEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMELYDLNRDLLNGYKIRWI TELLGNLKAVNQAIQRAGRLRVGKPKNI ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	rvq
LIKKVQEAVLDK WVNDPHRMDRRLLAI AHASDVLENAFAPLLDEQYDLATKRVRI LDPEVECLKANTNEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMELYDLNRDLLNGYKIRWI TELLGNLKAVNQAIQRAGRLRVGKPKNG ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	KNL
AHASDVLENAFAPLLDEQYDLATKRVRI LDPEVECLKANTNEVLWAVVAAFTK 761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMELYDLNRDLLNGYKIRWI TELLGNLKAVNQAIQRAGRLRVGKPKNG ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	
761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMELYDLNRDLLNGYKIRWI TELLGNLKAVNQAIQRAGRLRVGKPKNO ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	
761 2111 A 6078 833 390 IVSFHLSGFKKFVRPFSFLSVHGLQVDEY HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMELYDLNRDLLNGYKIRWI TELLGNLKAVNQAIQRAGRLRVGKPKNI ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	ALLD
HQKLSADMADHSNLIRSLLVGAEDARLI MKTMKSRYMELYDLNRDLLNGYKIRWI TELLGNLKAVNQAIQRAGRLRVGKPKNI ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	HOV
MKTMKSRYMELYDLNRDLLNGYKIRWI TELLGNLKAVNQAIQRAGRLRVGKPKNG ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	
TELLGNLKAVNQAIQRAGRLRVGKPKNO ACRDAIRSNNINTLFKIMRVGTASS 762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQT SFGINSNNQLAEKVRLRLRYEEAKRRIAN	
762 2112 A 6079 2 2686 KKAITCGEKEKQDLIKSLAMLKDGFRTD HSDLWSSSSSLESSSFPLPKQYLDVSSQT SFGINSNNQLAEKVRLRLRYEEAKRRIAN	VIT
HSDLWSSSSSLESSSFPLPKQYLDVSSQTI SFGINSNNQLAEKVRLRLRYEEAKRRIAN	`
SFGINSNNQLAEKVRLRLRYEEAKRRIAN	
EMRFISPRKWTQGEVEQLEMARKRLEKI AARDTQSKALTERLKLNSKRNQLVRELE	
RQVATLHSQLKSLSSSMQSLSSGSSPGSL	
GSLVASSLDSSTSASFTDLYYDPFEQLDS	
SKVEFLLLEGATGFRPSGCITTIHEDEVA	
KAEGGGRLQALRSLSGTPKSMTSLSPRS	
PSPPCSPLMADPLLAGDAFLNSLEFEDPE	LSA
TLCELSLGNSAQERYRLEEPGTEGKQLGG	
NTAQGCGLKVACVSAAVSDESVAGDSG	
ASVQRLGASEAAAFDSDESEAVGATRIQ	
YDEKNKQFAILIIQLSNLSALLQQQDQKV VAVLPCSESTTCLFRTRPLDASDTLVFNE	NIK
VSMSYPALHQKTLRVDVCTTDRSHLEEC	
AQISLAEVCRSGERSTRWYNLLSYKYLK	
RELKPVGVMAPASGPASTDAVSALLEQT	AVE
LEKRQEGRSSTQTLEDSWRYEETSENEA	
EEEEEVEEEEGEEDVFTEKASPDMDGYPA	
VDKETNTETPAPSPTVVRPKDRRVGTPSC	
LRGSTIIRSKTFSPGPQSQYVCRLNRSDSD	
LSKKPPFVRNSLERRSVRMKRPSPPPQPS:	
SLRSERLIRTSLDLELDLQATRTWHSQLT	
VLKELKEQLEQAKSHGEKELPQWLREDE LLLRMLEKRMDRAEHMGELQTDKMMR.	
. KDVHRLRGQSCKEPPEVQSFREKMAFFT	
MNIPALSADDV	AER I
763 2113 A 6082 3 1558 PHPIRFSKLCVSFNNQEYNQFCVIEEASKA	
VLENLTQGKMCLVPGKTRKLLFKFVAKT	- 1
VGKKIEITSVDLALGNETGRCVVLNWQG	NE
DAASSQEALQAARSFKRRPKLPDNEVHW	NE ED
IQASTMIISRVPNISVHLLHEPPALTNEMY	NE ED GGG GSII

البياة الأساة الأسد الهيدة التراسيد السداليد

SEQ ID NO: of nucl- eotide seq-	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496	Predicted beginning nucleotide location correspondi	Predicted end nucleotide location corresponding to last amino	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid residue of peptide sequence	acid residue of peptide sequence	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
764	2114	A	6093		1422	VTVQSHEKTQIRDVKLTAGLKPGQDANLTQK THVTLHGTELCDESYPALLTDIPVGDLHPGEQ LEKMLYVRCGTVGSRMFLVYVSYLINTTVEE KEIVCKCHKDETVTIETVFPFDVAVKFVSTKF EHLERVYADIPFLLMTDLLSASPWALTIVSSE LHLAPSMTTVDQLESQVDNVILQTGESASECF CLQCPSLGNIEGGVATGHYIISWKRTSAMENI PIITTVITLPHVIVENIPLHVNADLPSFGRVRES LPVKYIILQNKTDLVQDVEISVEPSDAFMFSG LKQIRLRILPGTEQEMLYNFYPLMAGYQQLPS LNINLLRFPNFTNQLLRRFIPTSIFVKPQGRLM DDTSIAAA AAADLANSNAGAAVGRKAGPRSPPSAPAPAP
						PPPAPAPPTLGNNHQESPGWRCCRPTLRERN ALMFNNELMADVHFVVGPPGATRTVPAHKY VLAVGSSVFYAMFYGDLAEVKSEIHIPDVEPA AFLILLKYMYSDEIDLEADTVLATLYAAKKYI VPALAKACVNFLETSLEAKNACVLLSQSRLF EEPELTQRCWEVIDAQAEMALRSEGFCEIDR QTLEIIVTREALNTKEAVVFEAVLNWAEAEC KRQGLPITPRNKRHVLGRALYLVRIPTMTLEE FANGAAQSDILTLEETHSIFLWYTATNKPRLD FPLTKRKGLAPQRCHRFQSSAYRSNQWRYRG RCDSIQFAVDRRVFIAGLGLYGSSSGKAEYSV KIELKRLGVVLAQNLTKFMSDGSSNTFPVWF EHPVQVEQDTFYTASAVLDGSELSYFGQEGM TEVQCGKVAFQFQCSSDSTNGTGVQGGQIPE LIFYA
765		Α .	6099	1-		SGFTHYAIYDFIVKGSCFCNVHADQCIPVHGF RPVKAPGTFHMVHGKCMCKHNTAGSHCQH CAPLYNDRPWEAADGKTGAPNECRTCKCNG HADTCHFDVNVWEASGNRSGGVCDDCQHN TEGQYCQRCKPGFYRDLRRPFSAPDACKPCS CHPVGSAVLPANSVTFCDPSNGDCPCKPGVA GRRCDRCMVGYWGFGDYGCRPCDCAGSCD PITGDCISSHTDIDWYHEVPDFRPVHNKSEPP WEWEDAQGFSALLHSGKCECKEQTLGNAKA FCGMKYSYVLKIKILSAHDKGTHVEVNVKIK KVLKSTKLKIFRGKRTLYPESWTDRGCTCPIL NPGLEYLVAGHEDIRTGKLIVNMKSFVQHWK PSLGRKVMDILKRECK
766	2116	A	6103	2	384	MTAAATATVLKEGVLEKRSGGLLQLWKRKR CVLTERGLQLFEAKGTGGRPKELSFARIKAVE CVESTGRHIYFTLVTEGGGEIDFRCPLEDPGW NAQITLGLVKFKNQQAIQTVRARQSLGTGTL VS
767	2117	A	6106	1	542	SGSSHASDGSGFQELRICSEDQTPLIAGMCSLP MARYYIIKYADQKALYTRDGQLLVGDPVAD NCCAEKICTLPNRGLDRTKVPIFLGIQGGSRC LACVETEEGPSLQLEDVNIEELYKGGEEATRF TFFQSSSGSAFRLEAAAWPGWFLCGPAEPQQ PVQLTKESEPSARTKFYFEQSW
768	2118	A	6109	3	292	FILQAVLQLSSQEARYKAFGTCVSHIGAILAF YTPSVISSVMHRVARCAAPHVHILLANFYLLF PPMVNPIIYGVKTKQIRDSLGSIPEKGCVNRE
769	2119	A	6110	1	711	RHEPSCSNGVASTKSKQNHSKYPAPSSSSSSS SSSSSSSSSSVNYSESNSTDSTKSQHHSSTSNQ ETSDSEMEMEAEHYPNGVLGSMSTRIVNGAY KHEDLQTDESSMDDRHPRRQLCGGNQAATE

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	J	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	i	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	l.	1		peptide	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
		1		sequence	1	nucleotide insertion
	 	1	 	sequence		RILFGRELQALSEQLGREYGKNLAHTEMLQD
	ļ	ļ		ļ		AFSLLAYSDPWSCPVGQQLDPIQREPVCAAL
	ļ	1				NSAILESONLPKOPPLMLALGOASECLRLMA
		<u> </u>				RAGLGSCSFARVDDYLH
770	2120	Ā	6125	2	570	YFGLNLHVQHLGNNVFLLQTI.FGAVILLANC
		1			[VAPWALKYMNRRASQMLLMFLLAICLLAIIF
						VPQEMQMLREVLATLGLGASALANTLAFAH
			ľ			GNEVIPTIIRARAMGINATFANIAGALAPLMM
1		ł				ILSVYSPPLPWIIYGVFPFISGFAFLLLPETRNK
771	2121	A	6126	909	353	PLFDTIQDEKNERKDPREPKQEDPRVEVTQF
'''	2121	^	0120	909	درد ا	RSFVLDTASAICNYNAHYKNHPKYWCRGYF RDYCNIIAFSPNSTNHVALRDTGNQLIVTMSC
		ļ]]	LTKEDTGWYWCGIQRDFARDDMDFTELIVT
						DDKGTLANDFWSGKDLSGNKTRSCKAPKVV
	l	1	1			RKADRSRTSILIICILITGLGIISVISHLTKRRRS
	1	1	1			QRNRRVGNTLKPFSRVLTPKEMAPTEQM
772	2122	Α	6148	7	810	FVLGILALSHTISPFMNKFFPASFPNRQYQLLF
						TQGSGENKEEIINYEFDTKDLVCLGLSSIVGV
1		l				WYLLRKHWIANNLFGLAFSLNGVELLHLNN
						VSTGCILLGGLFIYDVFWVFGTNVMVTVAKS
						FEAPIKLVFPQDLLEKGLEANNFAMLGLGDV
j		ĺ)			VIPGIFIALLERFDISLKKNTHTYFYTSFAAYIF
	1	1				GLGLTIFIMHIFKHAQPALLYLVPACIGFPVLV ALAKGEVTEMFSYEESNPKDPAAVTESKEGT
	l					EASASKGLEKKEK
773	2123	A	6161	3	1088	COPMLYTRKNHPKLLLRRTESVAEKMLTNW
		i		•		FTFLLYKFLKESAGEPLFMLYCAIKHQMEKG
Ì		ļ				PIDAITGEARYSLSEDKLIRHLIDYKTLTLNCV
1		ļ		•		NPENENAPEVPVKGLDCDTGTQAKEKLLDA
	1	j				AYKGVPYSQRPKAADMDLEWRQGRMARJIL
						QDEDVTTKIDNDWKRLNTLAHYQVTDGSSV
						ALVPKQTSAYNISNSSTFTKSLSRYESMLRTA
1	· ·					SSPDSLRSRTPMITPDLESGTKLWHLVKNHDH
ĺ	[LDQREGDRGSKMVSEIYLTRLLATKGTLQKF VDDLFETIFSTAHRGSALPLAIKYMFDFLDEO
	1					ADKHQIHDADVRHTWKSNCLPLRFWVNVIK
						NPQFVFDIHKNSITDACLSVV
774	2124	A	6163	860	125	KTAVKKRNLNPVFNETLRYSVPQAELQGRVL
						SLSVWHRESLGRNIFLGEVEVPLDTWDWGSE
1		1				PTWLPLQPRVPPSPDDLPSRGLLALSLKYVPA
						GSEGAGLPPSGELHFWVKEARDLLPLRAGSL
						DTYVQCFVLPDDSRASRQRTRVVRRSLSPVF
						NHTMVYDGFGPADLRQACAELSLWDHGALA
	-			-		NRQLGGTRLSLGTGSSYGLQVPWMDSTPEEK
775	2125	A	6191	2	392	QI.WQALLEQPCEWVDGLLPLRTNLAPRT ARGIGSLGRDHSGSGGGTGMAGAWVRKAAD
'''	س.ت	1	0171	~	J72	YVRSKDFRDYLMSTHFWGPVANWGLPIAAIT
		'			1	DMK\KSPEIISRRMTFAL*CYSLTTVRFAHYVO
						PWNWLMLGCHTAVDFDQLISSMPCISHGMT
						ASASAL
776	2126	A	6217	1	827	FRGYWGVREAFTDASWSGGLGPGKPGMKIT
						RQKHAKKHLGFFRNNFGVREPYQILLDGTFC
]				ĺ	ļ	QAALRGRIQLREQLPRYLMGETQLCTTRCVL
						KELETLGKDLYGAKLIAQKCQVRNCPHFKNA
1			İ	ŀ		VSGSECLLSMVEEGNPHHYFVATQDQNLSVK
1				ļ	ļ	VKKKPGVPLMFIIQNTMVLDKPSPKTIAFVKA
				ļ		VESG\RLSQCMRKKVSNISKRNRV**KTLNRG
					į	RRKKRKKISGPNPLSCLKKKKKAPDTQSSASE
L		L		,		KKRKRKRIRNRSNPKVLSEKQNAEGE

OFO IF	L CEOTO	1 1/4	1000	I B. 31.4.3	1 5 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 	
SEQ ID		Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	NO: of peptide	BOO	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-		in USSN	location	location	F=Phenylalanine, G=Glycine, H=Histidine,
seq-	neuce		09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine,
uence	donce		914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine,
uchec	1	1	1 314	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
į	1			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1		k	peptide	Sequence	/=possible nucleotide deletion, \=possible
ļ			ĺ	sequence		nucleotide insertion
777	2127	A	6236	1038	1402	YYQISSLPSIVGNGIFLWLLICIFLAKQGGSRL*
]	1					FQPFGRPRGGGHLRSGVLGQPGQHGETP/SFF
						YNSKISPALWGPPVIPSALGGEAGKSL*PRRQ
				ļ		RFQRGGIAPLPSRVRGRAKLFLKKK
778	2128	A	6237	422	913	ASFFHHHRGAFLLLLAIPGS*GQDQSLIHWSN
1	1	ł	1	ł	1	AVSNAD\LLDLK\N*LDH\LEEKMPL\EVKVVP
		1				PQVL\SEPN*RSGGCFSAPSFEVPPWTGEVKP/
1		i				SPQRDGGALG\QGPLGIPSDSILALLKKQT*RA
	1		ļ. :			LLNWPLGSLRRSSCFGGQDGQDLKPRSGLGC
			ľ			NSFRYRR
779	2129	A	6249	420	36	ARAPSPSFSVRDVELSDPARERGEMPVAVGP
1	1	1				YGQSQPSCFDRVKMGFVMGCAVGMAAGAL
		1				FGTFSCLSSILVSSSG/SGMRGRELMGGIGKTM
				•		MQSGGTFGTFMAIGMGIRC*PWLPTTSVPSH
L						QSQPMY
780	2130	Α	6263	415	1380	RIMRMCDRGIQMLITTVGAFAAFSLMTIAVG
1]			TDYWLYSRGVCRTKSTSDNETSRKNEEVMT
	İ					HSGLWRTCCLEGAFRGVCKKIDHFPEDADYE
	Ì					QDTAEYLLRAVRASSVFPILSVTLLFFGGLCV
1	1	l				AASEFHRSRHNVILSAGIFFVSAGLSNIIGIIVYI
1	1					S\ANAGRTPGQR\DSKKSYSYGWSF/YFSGAFS
						FIIGR/IIC*GVGLPWHIYIEKHQQLRAKSHSEF
						LKKSTFARLPPYRYRFRRRSSSRSTEPRSRDLS
	l l					PISKGFHTIPSTDISMFTLSRDPSKITMGTLLNS
J						DRDHAFLQFHNSTPKEFKESLHNNPANRRTT
						PV
781	2131	A	6274	832	318	RIIKVKDLKQTLAIKTAYPRCKCLVEMDQIFH
1		l.				LQVKQKQLACLCTWQARDPDCPPSTKVVL/L
1	J]	•		VGPGMGCMVALFQDSIAWSNKSMPSSLSAIS
			1			QSPCQVQAPEGPSSFHLPTLSFTTCLSWQGGD
			1			LEFLGDLKGCSELKNFQELITQSALVHPKADV
						WWYCGRPLLGTLPSN
782	2132	A	6281	1324	393	WISLPSSLLCRKNGSSAEDDRR\GEPSAEEAEG
1						EREDWGIGSA*SVGAVSKVPSARF*RTYPS\E
1	1					DEEEVTHQKSSSSDSNSEEHRKKKTSRSRNK
					.	KKRKNKSSKRKHRKYSDSDSNSESDTNSDSD
						DDKKRVKAKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKKK
			<u> </u>			ESSDSSCKDSEEDLSEATWMEQPNVADTMDL
1	1	1		i	·	IGPEAPIIHTSQDEKPLKYGHALLPGEGAAMA
1						EYVKAGKRIPRRGEIGLTSEEIGSFECSGYVM
1					İ	SGSRHRRMEAVRLRKENQIYSADEKRALASF
783	2133	A	6305	201	1032	NQEERRKRESKILASFREMVHKKTKGKDDK
1,02	2133	Λ.	0303	201	1032	WDDYPQGALRREAAEGLHFLGPPGRVRGQ
			-	j		LRGITGPAWYCHSPSHSLLSAFCHLPTPSRCP
1						AMARPPVPGSVVVPNWHES/RRGQGVPGLHS
				ļ		AQEPPAGVWAA*AASAAAA\LSIDTASYKIFV
			[SGKSGVGKTALVAKLAGLEVPVVHHETTGIQ
1] }			TTVVFWPAKLQASSRVVMFRFEFWDCGESA
					Í	LKKFDHMLLACMENTDAFLFLFSFTDRASFE
ŀ						DLPGQLARIAGEAPGVVRMVIGSKFDQYMHT
784	2134	A	6308	06	<u> </u>	DVPERDLTAFRQAWELPLLRVKSVPGRRLG
104	4134	^	9000	86	96	GSSPDPASLITMKNQDKKNGAAKQSNPKSSP
	1 1			j		GQPEAGPEGAQERPSQAAPAVEAEGPGSSQA
1					1	PRKPEGAQARTAQSGALRDVSEELSRQLEDIL
						STYCVDNNQGGPGEDGAQGEPAEPEDAEKSR
]					TYVARNGEPEPTPVVNGEKEPSKGDPNTEEIR
1						QSDEVGDRDHRRPQEKKKAKGLGKEITLLM
1					1	QTLNTLSTPEEKLAALCKKYAELLEEHRNSQ
Ь			L			KQMKLLQKKQSQLVQEKDHLRGEHSKAVLA

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino said some (A. Alaria C. C.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	denoc		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1		1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1				peptide	sequence	/=possible nucleotide deletion, \=possible
1	ļ	1		sequence	ļ	nucleotide insertion
	 	 	 	Sequence	ļ	RSKLESLCRELQRHNRSLKEEGVQRAREEEE
1			1			VDVEVTOIEOVTI NOIOI OMBOIDIEDNISM B
1		ł	1	}		KRKEVTSHFQVTLNDIQLQMEQHNERNSKLR
ì	ľ	i	l	į	1	QENMELAERLKKLIEQYELREEHIDKVFKHK
1			ļ	<u> </u>		DLQQQLVDAKLQQAQEMLKEAEERHQREKD
					1	FLLKEAVESQRMCELMKQQETHLKQQLALY
1		1		ļ	1	TEKFEEFONTLSKSSEVFTTFKQEMEKMTKKI
						KKLEKETIMYRSRWESSNKALLEMAEEKTV
						RDKELEGLQVKIQRLEKLCRALQT/GAQ*PVR
785	2126		(210	1.402	000	GQRWGSHRTSAVRIFS
/65	2135	Ā	6319	1493	889	SPQGPLLRSVSPVSAGASSVTPGGAQPGVTTT
ł		ļ			ŀ	PPSLVAVAPAPGSAAGPAAGWQ*HAGCR/WT
			ļ ,			KLPWSWGMRPMKIFFSEEYRSISTRISHDAL*
					i	EKCTQPAKPLSMIR\TGSSVSPG/PLVKWNWT
						RREFRNSGTRVVSSCCGMSCMYSFLGHCSV/S
1		Ì	1 1		Ì	QDLPLVHVDVGWQPPLGPTVGLRPGLLPLHD
						TTPCQKLVVDDLDWA
786	2136	A	6320	551	135	RWLPVAECDSSCVGCTGEGPGNCKECISGYA
		l	,			REHGQCADVDECSLAEKTCVRKNENCYNTP
		ļ				GSYVCVCPDGFEET/RRCLCAAGRG*SHRRRK
1	1	[[PDTAALPRRPVMCRTYPLNYSEGCPVENVAL
						RMPSPAVDSGGERLPAL
787	2137	Α	6330	1693	227	DYVLTAELHRQRSPGVSFGLSVFNLMNAIMG
1						SGILGLAYVMANTGVFGFSFLLLTVALLASYS
1						VHLLLSMCIQTAYLGP*TNYFMVLPAH*LTCL
1				,		PLIEFLQSL*NSL*AVTSYEDLGLFAFGLPGKL
]						VVAGTIIIQNIGAMSSYLLIIKTELPAAIAEFLT
			§	•		GDYSRYWYLDGQTLLIIICVGIVFPLALLPKIG
						FLGYTSSLSFFFMMFFALVVIIKKWSIPCPLTL
						NYVEKGFQISNVTDDCKPKLFHFSKESAYALP
						TMAFSFLCHTSILPIYCELQSPSKKRMQNVTN
			j			TAIALSFLIYFISALFGYLTFYD/GTTKAQRGE
1		ľ				VTCHRIKDKVESELLKG***IP*SHDVVVMT\V
						KLCILFAVLL\TVPLIHFPARKAVTMMFFSNFP
j						FSWIRHFLITLALNIIIVLLAIYVPDIRNVFGVV
[ľ		ľ	GASTSTCLIFIFPGLFYLKLSREDFLSWKKLGV
					l	GCFC/LLSFKTSILRNSLSVYIILPASRKSIYFKI
788	2138	A	6351	1	6622	PRSLCFSLWAEAAVLADGGLRRRRRLLRGTM
			ŀ			SASFVPNGASLEDCHCNLFCLADLTGIKWKK
						YVWQGPTSAPILFPVTEEDPILSSFSRCLKADV
i 1		ľ	i 1		İ	LG/VWRRDQRPERRE\L*IFWGGEDP\VLLTLF
						TMTYQKKKMECGRMDFPMNAVLCFSKAVH
		į		1		NLLERCLMNRNFVRIGKWFVKPYEKDEKPIN
}	1	- 1	!	}		KSEHLSCSFTFFLHGDSNVCTSVEINQHQPVY
		1		İ		LLSEEHITLAQQSNSPFQVILCPFGLNGTLTGQ
	l	- 1				AFKMSDSATKKLIGEWKQFYPISCCLKEMSE
	į	1	1		ŀ	EKQEDMDWEDDSLAAVEVLVAGVRMIYPAC
	ľ		l	1		
1 1	}	Į	- 1	1	ļ	FVLVPQSDIPTPSPVGSTHCSSSCLGVHQVPAS
	j		. [İ	1	TRDPAMSSVTLTPPTSPEEVQTVDPQSVQKW
				Ţ	[VKFSSVSDGFNSDSTSHHGGKIPRKLANHVV
	ł	İ		l		DRVWQECNMNRAQNKRKYSASSGGLCEEAT
	[ı	- 1	Ī	l	AAKVASWDFVEATQRTNCSCLRHKNLKSRN
j . l	!		- 1		į	AGQQGQAPSLGQQQQILPKHKTNEKQEKSEK
			1	l	j	PQKRPLTPFHHRVSVSDDVGMD\ADS\ASQRL
İ	ľ		. [ŀ]	V\ISAP\DSQ\VRFSNIR\TNDVAK\TPQMHGTE
	ł		İ	į	i	MANSPQPPPLSP\HPCDVVDEGVTKTPSTPQS
[ſ	į	f	ĺ	QHFYQMPTPDPLVPSKPMEDRIDSLSQSFPPQ
		}	į	-	l	YQEAVEPTVYVGTAVNLEEDEANIAWKYYK
	1	ı	į	1	l	FPKKKDVEFLPPQLPSDKFKDDPVGPFGQESV
	- 1	ı	Ì	j	· · · · · · · · · · · · · · · · · · ·	TSVTELMVQCKKPLKVSDELVQQYQIKNQCL
	1		L			SAIASDAEQEPKIDPYAFVEGDEEFLFPDKKD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
						RQNSEREAGKKHK VEDGTSSVTVLSHEEDA MSLFSPSIKQDAPRPTSHARPPSTSLIYDSDLA VSYTDLDNLFNSDEDELTPGSKRSANGSDDK ASCKESKTGNLDPLSCISTADLHKMYPTPPSL EQHIMGFSPMNMNNKEYGSMDTTPGGTVLE GNSSSIGAQFKIEVDEGFCSPKPSEIKDFSYVY KPENCQILVGCSMFAPLKILPSQYLPLIKLPEE CIYRQSWTVGKLELLSSGPSMPFIKEGDGSNM DQEYGTAYTPQTHTSCGMPPSSAPPSNSGAGI LPSPSTPRFPTPRTPRTPRTPRTPRGAGGPASAQGS VKYENSDLYSPASTPSTCRPLNSVEPATVPSIP EAHSLYVNLILSESVMNLFKDCNSDSCCICVC NMNIKGADVGVYIPDPTQEAQYRCTCGFSAV MNRKFGNNSGLFFEDELDIIGRNTDCGKEAE KRFEALRATSAEHVNGGLKESEKLSDDLILLL QDQCTNLFSPFGAADQDPFPKSGVISNWVRV EERDCCNDCYLALEHGRQFMDNMSGGKVDE ALVKSSCLHPWSKRNDVSMQCSQDILRMLLS LQPVLQDAIQKKRTVRPWGVQGPLTWQQFH KMAGRGSYGTDESPEPLPIPTFLLGYDYDYLV LSPFALPYWERLMLEPYGSQRDIAYVVLCPE NEALLNGAKSFFRDLTAIYESCRLGQHRPVSR LLTDGIMRVGSTASKKLSEKLVAEWFSQAAD GNNEAFSKLKLYAQVCRYDLGPYLASLPLDS SLLSQPNLVAPTSQSLITPPQMTNTGNANTPS ATLASAASSTMTVTSGVAISTSVATANSTLTT ASTSSSSSSNLNSGVSSNKLPSPPPFGSMNSNA AGSMSTQANTVQSGQLGGQQTSALQTAGISG ESSSLPTQPHPDVSESTMDRDKVGIPTDGDSH AVTYPPAIVVYIIDPFTYENTDESTNSSSVWTL GLLRCFLEMVQTLPPHIKSTVSVQIIPCQYLLQ PVKHEDREIYPQHLKSLAFSAFTQCRRPLPTS TNVKTLTGFGPGLAMETALRSPDRPECIRLYA PPFILAPVKDKQTELGETFGEAGQKYNVLFV GYCLSHDQRWILASCTDLYGELETCIINIDVP
					-	NRARRKKSSARKFGLQKLWEWCLGLVQMSS LPWRVVIGRLGRIGHGELKDWSCLLSRRNLQ SLSKRLKDMCRMCGISAADSPSILSACLVAM EPQGSFVIMPDSVSTGSVFGRSTTLNMQTSQL NTPQDTSCTHILVFPTSASVQVASATYTTENL DLAFNENNDGADGMGIFDLLDTGDDLDPDII NILPASPTGSPVHSPGSHYPHGGDAGKGQSTD RLLSTEPHEEVPNILQQPLALGYFVSTAKAGP LPDWFWSACPQAQYQCPLFLKASLHLHVPSV QSDELLHSKHSHPLDSNQTSDVLRFVLEQYN ALSWLTCDPATQDRRSCLPIHFVVLNQLYNFI MNML
789	2139	A	6359	1	2002	TGTTTEDGLDVMGVVPLKGQAFLPLVPEPRR LPVGPLLRALATCHALSRLQDTPVGDPMDLK MVESTGWVLEEEPAADSAFGTQVLAVMRPP LWEPQLQAMEEPPVPVSVLHRFPFSSALQRM SVVVAWPGATQPEAYVKGSPELVAGLCNPET VPTDFAQMLQSYTAAGYRVVALASKPLPSVP SLEAAQQLTRDTVEGDLSLLGLLVMRNLLKP QTTPVIQALRRTRIRAVMVTGDNLQTAVTVA RGCGMVAPQEHLIIVHATHPEGQPASLEFLP MESPTAVNGVKDPDQAASYTVEPDPRSRHLA LSGPTFGIIVKHFPKLLPKVLVQGTVFARMAP EQKTELVCELQKLQYCVGMCGDGANDCGAL KAADVGISLSQAEASVVSPFTSSMASIECVPM

Sol Discourse Sol	CEC ID	I CEO M	Mat	LODO	Predicted	Deadlet - 1 1	Amino neid gogueras (Amilianias Caro
nuclocide coefficies of the contract of the co	SEQ ID	SEQ ID	Met	SEQ		Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
coide seq- uence sol- 9/496 periodic sequence corresponding bist amino acid residue of peptide sequence i-losoleucine, K-l-ysine, L-Leucine, 0-Guttamine, R-Arginine, S-Serine, 1-Threonine, V-waline, W-Thytophan, 7-Thryosine, X-Usine, W-Thytophan, 7-Thryosine, X-Usine, W-Thytophan, 7-Thryosine, X-Usine, W-Thytophan, 7-Thryosine, X-Usine, W-Thytophan, 7-Thryosine, X-Usine, W-Thytophan, 7-Thryosine, X-Usine, W-Thytophan, 7-Thryosine, X-Usine, W-Thytophan, 7-Thryosine, X-Usine, W-Thytophan, 7-Thryosine, X-Usine, W-Thytophan, 7-Thryosine, X-Usine, W-Thytophan, 7-Thryosine, X-Usine, W-Thytophan, 7-Thryosine, X-Usine, W-Thytophan, 7-Thryosine, X-Usine, W-Thytophan, 7-Thryosine, X-Usine, W-Thryosine, X-Usine, W-Thryosine, X-Usine, W-Thryosine, X-Usine, W-Thryosine, X-Usine, W-Thryosine, X-Usine, W-Thryosine, X-Usine, W-Thryosine, X-Usine, W-Thryosine, X-Usine, W-Thryosine, X-Usine, W-Thryosine, X-Usine, W-Thryosine, X-Usine, W-Thryosine, X-Usine, W-Thryosine, X-Usine, W-Thryosine, X-Us			noa	1		1	
Sequence		1		1	1	1 '	
uence 914 ng to first mino acid residue of peptide sequence pept							
mino acid residue of peptide requence per per per per per per per per per pe	seq-	uence	(1			
Persidue of peptide sequence	uence		1	914		acid residue	Q=Glutamine, R=Arginine, S=Serine,
residue of peptide pep			1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ppptide			l		residue of		Y=Tyrosine, X=Unknown, *=Stop codon.
	1.	1			pentide	•	
VIREGECSLDTSENVEKYMALYSLTQFESVLL VINTHLOGLOFLANDLYTTTVALVAMSRTGP ALVIGRVEPPGALLSVPVISSLLLQMVLVTG VOLGGYFLILAOPWYLSRTCHAPDLEPNY FOLYOFILLSALTSSVLVVLVLSGLHOFLALR NYFFLLASALTSSVLVVLVLSGLHOFLALR NYFFLLASALTSSVLVVLVLSGLHOFLALR NTTOTGELLLVGLVTLNEVOGLHAGEARR VPPRLPAPPAQAGISKRRRKQLERELAEOPW PPLPAGPER 790 2140 A 6380 76 1059 SSAGSARKLQVMALAARLWRLLPFRRGAP FOLYOFILSALTSSKRGHCGPCFRGGEFWGNOFG FERGELLLSALSYLGFETYQVISQAAVVHATA KVEELEQADVLYSEGFEKKI,ULVY EALEYJKARRATLEKNEGTEKKI,ULVY EALEYJKARRATLEKNEGSFASHKWYANCLSDV GOYEGIKAKIANAVIKEHFERAJELNPKDATS HILMGINCYTTEAMPWYGRRIANACLQP **PPYFEKALGYPHYAGRAPHYORIANACLQP **PPYFEKALGYPHYYAGRAPHYOLTHIANACLQP **PPYFEKALGYPHYAGRAPHYORIANACLQP **PPYFEKALGYPHYAGRAPHYORIANACLQP **PPYFEKALGYPHYAGRAPHYORIANACLQP **PPYFEKALGYPHYAGRAPHYORIANACLQP **PPYFECALGYPHYAGRAPHYORIANACLQP **PPYFEKALGYPHYAGRAPHYORIANACLQP **PPYFEKALGYPHYAGRAPHYORIANACLQP **PPYFECALGYPHYAGRAPHYORIANACLQP **PPYFECALGYPHYAGRAPHYORIANACLQP **PPYFECALGYPHYAGRAPHYORIANACLQP **PPYFECALGYPHYAGRAPHYORIANACLQP **PPYFECALGYPHYAGRAPHYORIANACLQP **PPYFECALGYPHYAGRAPHYORIANACLAP **PPYFECALGYPHYAGRAPHYORIANACLQP **PPYFECALGYPHYAGRAPHYORIANACLAP **PPYFECALGYPHYAGRAPHYORIANACLAP **PPYFECALGYPHYAGRAPHYORIANACLAP **PPYFECALGYPHYAGRAPHYORIANACLAP **PPYFECALGYPHYAGRAPHYO							
THINTINGDLQFLAIDLUTITIVAVLMSRTGG ALVLGRYRPGALLSVPVISSLIQUMIVITG VQLGGYFLITLAQPWPVPLINITVAAPDNLERYI PETVVYSSSRQYILLAAAVSKQAFRRPLTIN NVPFLLASAL*SSVLVVLVLSPGLHGGFLAR HINTINGFKLLLVGLVLINIVOGLHAGEARR VPFRLPAPPAQAGSKKRRKQLERELAEQPW PPLPAGPIR 790 2140 A 6380 76 1059 SSAGSARKLQVMALAARLWRLIPFRRGAAP PPLPAGPIR ROBERT STANDAY STANDAY SAGSARKLQVMALAARLWRLIPFRRGAAP FRIGLILSALSYLOFETYQVISQAAVVHATA KVEEILEQADYLYESGETIKLYQLITQYKESE DAELLWRLARASRDVAQLSRTISEEKKLLVY EALEYARRA/LEKNESSTASHKWYACLSDV ODYEGIRAKIANA YILKHIPKEALELHVRLOSDV OYEGIRAKIANA YILKHIPKEALELHVRLOSDV OYEGIRAKIANA YILKHIPKEALELHVRLOSDV OYEGIRAAQLITSSESIN HIMGIWCTTAEMPWYQRRIANACLQUP **PPYYEKALGYHRAEQVDPN**SYKULLLG KYYLKLHINKKLAAFWLMKAKDYPAHTEED KQCTTAAQLITSSESIN 791 2141 A 6434 3 1460 IALLIVDGLAWDDGGGLALLHSPSKLLL*QDS SGMSYAWWCTITRAFKSLLCHUCYSIGPQ **VTCPQQDACKE*KSTAN**GG*RE**PQVLFF **PPSWAKGGMSKKKRQRDSLYARDYSIGPS VTCPQQDACKE*KSTAN**GG*RE**PQVLFF **ALSINA**VKFGMSKKKQRDSLYARDYSIGPS DQSGLDMTGIKQRQPSYDLTSVPNLFTYNSS PNNGQLAFGTYMTEDRIAQNIKTSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QCAQIQTHAIQYVVEPKARKITGFMELCQMDQ ILLLKSGCLEVVLVRNCRAFFPLNNTVLFEG KYGGMQMKALGSDDLVVEAFDPAKNLCSL QCATCHAIQHQVENHDDTLAKLIAKPITTA VCNLHGEKLQVFKQSHEEIVTTLFPPLYKEL **PPCATACK** 792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMLELTWAVILTFY LGILEALGWGPGA**CDMCTGEEKRELIPPA LGYGKGCKGKPPSTLIPHDLLEIRNGRSH ESPGEMDLADDWLSCHOMYGFTCHKTKGGDL MLYHYEGYLEKDGSLPHSTHKINNGQPIVPT LGILGALGWGPGA**CDMCTGEEKRELIPPA LGYGKGCKGKPPSTLIPHDLLEIRNGRSH ESPGENDLADDWLSKADEVALVKEFEKH GAVVMESHEDALVEDIPDKEDEDKDGFISA **ETYKHDBL** PSPCMDLADDWCLSKDEXALUKFEFKH GAVVMESHEDALVEDIPDKEDEDKDGFISA **ETYKHDBL** PSPCMDLADDWCLSKODEKSHTHKINNGQPIVPT LGILGALGWGPGA**CDMCTGEEKRERGIPPANN DLRSDDQDLTRAHHILDTEHPWDLISIPSEHH EATTCLEWDQGGFPGLISSWPTGGCWS MGVSTLANSWESSYGSLWEGGFHLWALS **WHINGYKLAHLVEKSCASSFGCKFRSRVFFS PSLILINGGAMEGWIAVTNSGLVTVSLQU** SQVLUSTISTESLCRLRARVALADJAFTGGGNI VVATADGSSASPVQFYKVCVSVVSECKCRIDT DLPSJFRRGTTIDHNSKSPATHLKFLARD MSGQVLLCASSGTSISVECWSKRSGCPYNN FQLSSVVYGNEQTILKKRILANTDLRAAVAVALADJAFTGGAN VVATADGSSASPVQFYKVCVSVVSECKCRIDT DLPSJFRRGTTIDHNSKSASS		 	 		boquanoo		
ALVLGRVRPFGALLSVPVLSSLLLQMVLYTG VQLGGVFLILAQPWEVLRRITVAAPDNLFNY ENTVVRSISSRQVILLAAAVSKGAPFRRPLIN NVPFLLASAL*SSVLVVLSSCOLHGPLALR NITDTGFKLLLVGLVTINFVGGLHAGERARP VPPLAAPPPAQAGISKKRFKQLERELAEQPW PPLAGPLR 790 2140 A 6380 76 1059 SSAGSARKLQVMALAARLWRLLFFRRGAAP SSAGSARKLQVMALAARLWRLLFFRRGAAP SSAGSARKLQVMALAARLWRLLFFRRGAAP SSAGSARKLQVMALAARLWRLLFFRRGAAP SRAGSARKLGVMALAARLWRLLFFRRGAAP GSRLPAGTISGSRGHCGPCFRRGEFWAGNOTG FKRGLLLSALSYLGFETVQVISQAAVVHATA KVEELEQADVLYSSESTASHKWYAICLSDV GDYEGIKAKIANAYIIKEHFEKAJELNFKDATS HILMGWCYTFAEMPWQRRIA*NACLQPP **PPYYEKALGYVHRARADYPFYTSKILLLG GTYTKLIKINKKLAAFWWQRIA*NACLQPP **PPYYEKALGYVHRARADYPPTAHTEDE KQIQTEAQQLTSFSEKN LYTIKLINKKLAAFWMAKADYPHATEDE KQIQTEAQQLTSFSEKN ALVINGGAARLARVYSSISSRGSISKILN NETSGTVANGSVDLPKSEGVYNNVSGQPSP DOSGLDMTGIKQRQFFDLTSVFNHATVSG PANGQLAPGTTMTEIDEIAQNIKSHLETCOY TMEELHQLAWQTHTYEEKAYQSKSREALW QCAQTHTMQTVVYEFAKATIGMELQNDQ ILLKSGCLEVVLVRACCAFPPLNNTVLFGG KYGGMWRALGSDLVRAEPPARINCSL QUTEESIALFSSAVLISPDRAWLEPRKVQKLQ REYFFALGHVIQKNHADDELAKLRIFTITA VCNLHGEKLQVFKQSHEEINATLEPPLYKELF PROCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFCFSNMRLFLWNAVLTLFV TSLGALARGWGGA'S CDMCYGEKRKLIPPA LGYGKEGKGKPPESTLIPNDLLEINGGPISH LGYGKEGKGKPPESTLIPNDLLEINGGPISH LGYGKEGKGKPPESTLIPNDLLEINGGPISH LGYGKEGKGKRPESTLIPNDLLEINGGPISH LGYGKEGKGKRPESTLIPNDLLEINGGPISH LGYGKEGKGKGRPESTLIPNDLLEINGGRSH LGYGKEGKGKGRPESTLIPNDLLEINGGBSH LGYGKEGKGKGRPESTLIPNDLLEINGGBSH LGYGKEGKGKGRPESTLIPNDLLEINGGBSH LGYGKEGKGKGRPESTLIPNDLLEINGGBSH LGYGKEGKGKGRPESTLIPNDLLEINGGBSH LGYGKEGKGKGRPESTLIPNDLLEINGGBSH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LGYGKEGKGKGRPESTLIPNDLLEINGSHH LG	1	1	ł		}	l	
VQLGGYFLITLAQPWFVPLNRIYAAPDNLPNY ENTVVFSISSPQYILAAAVSKGAPRRPLTN NVPFLLASAL'SSVLVVLVISPGLHGPLAIR NITDTGFKLLTUGU.TVLNFVGGHAGERAR VPPRLAPPRAQAGSKKRFKQLERELAEQPW PPLPAGPIR 790 2140 A 6380 76 1059 3SAGSARKLQVMALAARLWRLLPFRRGAAP GSRLPAGTSGSRGHCGPCERRGEWMGNEGT FKRGLLSALSYLGFETVQVISQAAVWHATA KVEELEQADVLYESGETEKLYQLLTQVKSES DAELLWRLARASRDVAQLSRTSEERKLLVY EALEYARRA/LEKNESSFASHKWYAICLSDV GYPGIKAKIAANYLIKEHFEKALELNFKDATS HILMGIWCYTFAEMPWYQRRIA*NACLQLPP *FPPYEKALGYTPHRAEQVOPNYSKNLLLG KTYLKLHNKKLAAFWLMKAKDYPAHTEED KQIOTEAQLLTSSEKN 791 2141 A 6434 3 1460 1ALLIVDELAWDDGGGLALLHSPSKLLI*ODS SMSTYWMCTITRAFKSLLCHLOYSIGPQ *VTCPGQDACKE*KSTAN*GG*RE**PQVLF AFLSNFAVKFGRMSKKQRDSI;YAEVQKHQQ RLGERGYQGSGEALARYYSSISSHGLSKLI NETSGTYANGSVIDLFKSEGYYNVVSGQPSP DOSGJDMTGKKQRQGYGEALARYYSSISSHGLSKLI NETSGTYANGSVIDLFKSEGYYNVVSGQPSP DOSGJDMTGKKQRQGYPTDLTSVTNLYTSS FNNGQLAPGITVMTEDRIAQNIKSHLETCQY TMEELHQLAWQTHTYEEKAYQSKSREALW QQCAIQTHAIQTVVEFAKRITGFMELCQNDQ LILLKSGCLEVVLYKNCRAFPPINNTVLFFG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QUTEELAFSSAVLISPGRAWLGFFTNTVLFFG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QUTEELAFSSAVLISPGRAWLGFTNTVLFFG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QUTEELAFSSAVLISPGRAWLGFTNTVLFFG TSLIGALIPSPAVLISPGRAWLFTNAVLTLFV TSLIGALIPSPAVLISPGRAWLFTNAVLTLFV TSLIGALIPSPAVLISPGRAWLFTNAVLTLFV TSLIGALIPSPAVLISPGRAWLFTNAVLTLFV TSLIGALIPSPAVLISPGRAWLFTNAVLTLFV TSLIGALIPSPAVLISPGRAWGHGRAWLDAVL LGYGKEGKGKRPFSTINNDLLEIRNGFSH LGYGKBGKGKGRPFSTINNDLLEIRNGFSH LGYGKBGKGKGKRPFSTINNDLLEIRNGFSH ESTGPUNDLNDWKLSKEDEKAYALKKEFEKH GAVVNESHEDALVERSGASSTGEKFSRIVZFS PRILTFIGGNAMEGWIAVTVSGLVTVSLQP SGQVLITSTESLCRLRPRAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDODLTRIMHIDTEPHPULISPISPEHH EATTCLEWDOSGFFGLFSRWYTSGLVTVSL WINNOWKLALLARVEKSGASSTGEKFSRIVZFS PSCLUTHNGGNAMEGWIAVTVSGLVTVSLQP SGQVLITSTESLCRLRRRVALDAIAFTGGGNI VVATADGSSASSPQTYKVCVSVVSEKGRIDP DILPSFRRRCTTIDLNRKKFPAITHLKFLARD MSGQVLLCASSGTSSTVECWSRKEGIPVANN FQUSPVYORGVQTILKKRILANTDLDRVA VALPRLIPSLTNTDLKVASSTOTGFYFGGLGLAL AFHOGSVFVORGVQTIKKRIANTDLDRVA		}				1	
PRITYVESISSPQ*ILIAAAVSKGAPFRRPLTIN NVPHLASAL*SSVLVVLVSCHCHGHALR NTIDTGFKLLLVGLVTINFVGGLHAGERARP VPPRLPAPPAQAGSKKRFKQLERELAEQPW PPLPAGPLR 790 2140 A 6380 76 1059 SSAGSARKLQVMALAARLWRLLPFRRGAAP SSAGSARKLQVMALAARLWRLLPFRRGAAP GSRLPAGTSGSRGHCGPCFRGGEPWGRNPGT FKRGLLLSALSYLGFETYQVISQAAVVHATA KVEELEQADVL*YSSGFASIKWTQTYGKESE DAELLWRLARASRDVAQLSRTSEERKLLLVY EALEPYAKRA/LEKNESFASIKWTACLSDV GDYEGIKAKIANAYIIKEHFEKAIELNFKDATS HILMGIWCYTFAEMPWQFRLA*NACLQ.PP **PPPYEKALGYPHRAEQVDPNFYSKNLLLLG KTYLKLHNKKLAAFWLMKAKDYPAHTEED KOQTTBAQULTSFSERN 791 2141 A 6434 3 1460 IALLIVDGLAWDDGGGLALLHISPSKLLL*ODS SGMSYVMVRCTITRAFKSLLCHLCQYSIGPQ **VTCPGQDACKE*KSTAN*GG*RE**PQVLFF AFISNPAVFFGMSKKGRDSLYABEVQHQQ RLQEQRQQQSGBAEALARYYSSISSINGLSNLN NTSGTYANGSVUDLFSEGGYYNVYGQFSP DOSGLDMTGIKQIKQEPYDLTSVPNLFTYNSS PNNGQLAFGITMTEDRIAQNIKSHLETQY TMEELHQLAWQTHTYEEIKAYQSKSRBALW QCCALQTHAIQYVYEFAKRITGFNELCQMDQ ILLKSGCLEVYLVRMCRAFNPLINTVLFEG KYGGMQMFKALGSDDLVRAAFDYSNLCHCQ EKYYFALGHVIQKNHLDDETLAKLIAKIPTITA VCNLHGERLQYSKGSHEVNTLFPYYKEL NPDCATACK 792 2142 A 6440 92 781 SSGTFRECTGNFFPCFSNMRLFLWNAVLTLFV TSLIGALIPSEPKIEVLQKFFICHRKTKGGDL MLYHFEGYLEKDGSLFISHKHNNOQPIWFT LGILRALKGWGGA*KDMCYGEKRKLIIPPA LGYGKEGKGKPPPSTLIPHDLEIRGRPSH ESPGEMDLADDWKLSKDEVKAYLKKEFEKH GAVYNESHHDALVERGPSH ESPGEMDLADDWKLSKDEVKAYLKKEFEKH GAVYNESHHDALVERGPSH ESPGEMDLADDWKLSKDEVKAYLKKEFEKH GAVYNESHHDALVERGPSH ESPGEMDLADDWKLSKDEVKAYLKKEFEKH GAVYNESHHDALVERGPSH ESPGEMDLADDWKLSKDEVKAYLKKEFEKH GAVYNESHHDALVERGPSH ESPGEMDLADDWKLSKDEVKAYLKKEFEKH GAVYNESHHDALVERGPSH ESPGEMDLADDWKLSKDEVKAYLKKEFEKH GAVYNESHHDALVERGPSH ESPGEMDLADDWKLSKDEVKAYLKKEFEKH GAVYNESHHDALVERGPSHEGRENT FYKHDBL 793 2143 A 6446 3201 152 PRIKRILVYTEEGGGARPEALGKIAPRTPLAG MGGRENGRAFERDHALVERGRSSHERGERFINGEP SPLTLRGGGAMBEGWIAVTYNGGLVYTSLLQP SGQVLITSTESLCRLRAVALADJAFTGGGNI VYATADGSAASSYGPTYKVCVSVVSEKCRIDT DILPSIFMCTTDLKWRINSHANDLDRNSA VALPKLIPISLTNTDLKVASDIGGYPGLGLAL AFHOGSVHVYRISLCIGHMAVYLSGLAVADDAFTGGGNI VYATADGSSAASPVQDCYKVCVSVVSEKCRIDT DILPSIFMCTTDLKWRINSHANDLDRNSA VALPKLIPISLTNTDLKVASDIGGYPGLGLALA AFHOGSVHVYRISLSLGTMAVYLSGLAVADDAFTGGGNI	1	ł		ł			
NVFFLLASAL*SSVLVVLVLSGLHAGEARR NITDTGKLLLVQLVTLNFVGGLAGEARR NITDTGKLLLVQLVTLNFVGGLAGEARR VPPR.PAPPAQAG\SKKRFKQLERELAEQPW PPLPAGPLR SAGSARKLQVMALAARLWRLLPFRRGAAP GSRLPAGTSGSRGHGGPCRRGFEWMGNPGF FRGLLSALSYLGFTYQVISQAAVVHATA KVEELEQADYLYESGETEKLYQLLTQYKESE DAELLWRLARASRDVAQLSRTSEEKKLLVY EALEYAKRAJZEKNESSSASHKWYALCLSDY GDYEGIGAKIANAYIKEHFEKARELDKDATS HILMGIWCYTFAEMPWYQRIA*NACLQLFQ FPPYEKALGYFHRAEQVDPHYSKSKLLLOG KTYLKLHNKILAAPWAMKAKDYPAHTEED KQIQTEAQALITSPEKNI. KTYLKLHNKILAAPWAMKAKDYPAHTEED KQIQTEAQALITSPEKNI. KTYLKLHNKILAAPWAMKAKDYPAHTEED KQIQTEAQALITSPEKNI. KTYLKLHNKILAAPWAMKAKDYPAHTEED KQIQTEAQALITSPEKNI. KTYLKLHNKILAAPWAMKAKDYPAHTEED KQIQTEAQALITSPEKNI. KTYLKLHNKILAAPWAMKAKDYPAHTEED KQIQTEAQALITSPEKNI. KTYLKLHNKILAAPWAMKAKDYPAHTEED KQIQTEAQALITSPEKNI. KTYLKLHNKILAAPWAMKAKDYPAHTEED KQIQTEAQALITSPEKNI. KTYLKLHNGAWAMI KTYLKLHNKILAAPWAMKAKDYPAHTEED KQIQTAAQALITSPEKNI. KTYLKLHNGAWAMI KTYLKLHNGAWAMI KTYLKLHNGAWAMI KTYLKLHNGAWAMI KTYLKLHNGAWAMI KTYLKLHNGAWAMI KTYLKLHNGAWAMI KTYLKLHNGAWAMI KYTKHAMI		ł	ļ	l			
NITDTGFKLLUGUNTINFVGGLHAGERARP VPPR.PAPPPAGGISKKRFQLERELAEQPW VPPR.PAPPPAGGISKKRFQLERELAEQPW VPPR.PAPPPAGGISKKRFQLERELAEQPW VPPR.PAPPPAGGISKKRFQLERELAEQPW VPRPRAGPLR SSAGSARKLQVMALAARLWRLLPFRRGAAP SSRLPAGTSGSKGHGGPCRRGFGFWGRNGFG FRRGLLISALSYLGFETYQVISQAAVVHATA KVEELEQADYLYSEGSTELALVGLTOYKESE DAELLWRLARASRDVAQLSRTSEEEKKLLVY EALEVARRAJLEKNESSSAGHKYALCISDV GDYEGIKAKIANAYIIKEHFEKAELINFKDAT HILMGIWCYTFAEMENYSGARTVALCISDV GDYEGIKAKIANAYIIKEHFEKAELINFKDAT HILMGIWCYTFAEMENYSGARTVALCISDV GOYEGIKAKIANAYIIKEHFEKAELINFKDAT HILMGIWCYTFAEMPYVQRRIA-NACLQLPP **PPPYEKALQIYFRRAEQVDPNFYSKNLLLLG KTYLKLHNKKLAAPULMKAADVPAHTEED KQIQTEAAQLITSPSEKN KYGLKLHAKADVJAKAADV			l	Ì			
VPPRLPAPPAGELR PPLAGPLR PPLAGPLR							NVPFLLASAL*SSVLVVLVLSPGLLHGPLALR
PPLPAGPLR	ŀ]		1	NITDTGFKLLLVGLVTLNFVGGLHAGERARP
PPLPAGPLR	1	1]		j	VPPRLPAPPPAQAG\SKKRFKQLERELAEOPW
790				i			
GSRLPAGTSGSRGHCGPCRERGFEWMCNPGT FKRGLLLSALSYLGFETYQVISQAAVVHATTA KVEEILEQADVLYPSGGFTEKLYQLITQYKESE DAELLWRLARASRDVAQLSRTSEEEKEKLLYV EALEYAKRA-LJCKENSESFASHKWYAICLSDV GDYEGIKAKIANAYIIKEHFEKAIELNFKDATS IHLMGIWCYTFAEMPWYGRRIAYNACLQLPP *PPPYEKALGYTHRAEQVDPNFYSKNLLLLG KTYLKLHNKKLAAFWLMKAKDYPAHTEED KQIQTEAQLLTSFSERN 791 2141 A 6434 3 1460 IALLIVDGLAWDDQGGGLALLHISPSKLIL-QDS SGMSYYMWRCHTTRAFFKSLCHICQYSIGPQ *VTCCPQQDACKE*KSTAN*GG*RE**PQVLFF AFLSNPAVKFGRMSKKQRDSLAVQKHQQ RQEQRQQQSGEAEALARVYSSSISNGLSNLN NETSGTYANGSVDLPKSESFLAVGVGKHQQ RQEQRQQSGEAEALARVYSSSISNGLSNLN NETSGTYANGSVDLPKSESFLAVGVGKHQQ QCAIQITHAIQYVVEFAKRTGFMELCONDQ ILLLKSGCLEVVLVMCKAFNPINNTVLFEG KYGGMQMFKALGSDDLVNBAFDFAKNLCSL QLTEEBIALFSSAVLISPDRAWLPSKYGELQ EKYYFALQHVIQKNHLDDETLAKLLAKIPTITA VCNLHGBKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFUNNAVLTLFV TSLIGALPFEPEVKIEVLQKFFICHKKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGIGLBALGWGPGA*KZMDFVGCHKKLTIPPA LGYGKEGKGKPPESTIJFNIDLLEIRNGGRSH ESFQEMDLNDDWKLSKDEVKAYLKKFEKH GAVNESHIDALVEDIPJKEDEDKDGFISAR FETYKHDBL 793 2143 A 6446 3201 152 PRIKRIVVTEEDGGARPEALGKIAPRTPAELG ARADGELVTALMCDLRRPAAGGMMDLAYV CEWEWSKSTHCPSVFLACAWSCRNLLAFTM DLRSDQQDLTRMIHLDTEHPWDLHSIPSEHH EAITCLEWDQGGFPGFISRWFTQGIKCWS MGVSTLANSWESSVGSLUVEGGFHL WALS. WLHNIGVKLALHYEKSGASSFGEKFSSKIKFS PSISTLFGGGMAMGGWIAATVSGLVTVSLLQVP SGQVLTSTESLCRLRARVALADIAFTGGGN1 VATADGSSASPVOPYKVCSVVSKEKCDID DLPSIEMRCTTDLNKDKFPAITHLKFLARD MSEQVILCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQFTILKWRILSATNDLDRVSS VALPKLIPISLTNTDLKVASDTQFYPGGLGLAL AFHDOSYHIVHRISLTJTMAVYSSAARPPYD 1	790	2140	A	6380	76	1059	
FKRGLLLSALSYLGFETYQVISQAAVVHATA KVEEILEQADVLYESGETEKLYOLTOYKESE DAELLWRLARASRDVAQLSRTSEEEKKLLVY EALEYKARA/LEKNESSFASHKWYAICLSDV GDYEGIKAKIANA YIKEHFEKAIELNPKDATS HILMGIWCYTFAEMPWYQRRIA*NACLQLPP **PPPYEKALGIYPHRAGQVDPHYYSKNLLLIG KTYLKLHNKKLAAFWLMKAKDYPAHTEED KQIQTEAAQLLTSFSEKN 791 2141 A 6434 3 1460 IALLIVDGILAWDDQGGLALHHISPSKLIL*QDS SGMSYVMVRCTITRAFKSLLCHICQYSIGPQ **VTCPGQDACKE*KSTAN*GG*RE**PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEGRQQSGEAEALARVSISNGLSNIAN NETSGTYANGSVIDLPKSEGYYNVVSQOPSP DQSGLDMTGJKQIKQEPIYDLTSYPNLTYTYSS FNNQLAPGITM*EDDRIAQNIKSSHLETCQY TMEELHQLAWQTHITYEEKAYQKSKSEALW QQCAIQTHAIQVVVEFAKITGFMELCONDQ QLTEESTALFSSAVLISPDRAWLEPKKVQKLQ EKIYFALQHVIQKNHLDDETLAKLJAKPITTA VCNLHGEKLQVPKQSHPEIVTLPPLYKELF NPDCATACK SKOTFRCCROPFFPCFSNMRLFLWNAVLTLFV TSLIGALJPEPEVKIEVLQKPFICHRKTKGGDL MLYHYBGYLEKDGSLFHSTHKHNNGQPIWFT LIGHEALKGWGGA*KDMCVGEKKKLJIPPA LGYGKSGKGKPPESTLIFNIDLLEIRNGFRSH ESGCEMDLNDDWKLSKGDEVAKKEFEKH GAVVNESHIDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRIKRIVYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEWWSKSTHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL BARDQOULTRIMHILDTEHPWDLHISISEHH EATCLEWDQSGFFGH-ISRNVKFS PSLTLFGGGMAMGGWIAVTYSGLVTVSLLQVP SGQVLTSTESLCRLRARVALADIAFTIGGGNI VATADGSSSASPVOPTKVCVSVVSECKDIDT DLPSLFMRCTTDLNRKDKFPATHLKFLARD MSEQVLLCASSQTSSIVEGWSLYKEGGHIALA FFIRGOSYHIVHRISLSLTMKVYSSSAAPRPVD VALADGSSVISITISLICHLARD AVAPKLIPISLTNTDLKVASDTQFYPGGLGLAL AFHDOSYHIVHRISLSLTMKVYSSSAAPRPVD	''"	2140	11	0560	'0	1037	
KVEEILEQADVI_VESGETEKL_VQLTOYKESE DABLLWELARARSDVAQLSEEEKLLVV EALEYAKRAADEKNESSFASHKWYAICLSDV GDYEGIKAKIANAYIIKEHFEKAIELINPKDATS HIHLMGIWCTIFAEMPWYGRRIA*NACLQLPP *PPPYEKALGYTHRAEQVDPNFYSKNILLI-G KTYLKLIHNKKLAFWIMKAKDYPAHTEED KQIQTEAQLLTSFSEKN 1460 IALLIVDGLAWDDQGGLALIHISPSKLIL-QDS SGMSYYMWRCTITRAFFKSLLCHICQYSIGPQ *VTICPGQDACKE*KSTIAN*GGRE*P9QUFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQSGEAEALARVYSSSISINGLSNIN NETSGTYANGSVDIJEKSESFVNYSSGOPSP DQSGLDMTIGIKQIKQEPIYDLTSVPNLFTYSS FNNGOLAPGTIMTEIDRIAQNIKSHLETCQY TIMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQITHAIQYVVEFAKRITGFMELCONDQ ILLLKSCCLEVVLVMKCRAFPILNNTVLFEG KYGGMQMFKALGSDDLVNEAFPFAKNICSL QEKIYFALQHVQKNHLDDETLAKLJAKPITTA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFUWNAVLILFY TSLIGALIPEEFVKIEVLQKFFCHRKTKGGDL MLYHYGGYLEKGOSLFHSTHKHNNGQPIWFT LGIELBALKGWGGA*KDMVCGKRKKLIPPA LGYGKRGKGKPPESTLIFNIDLLEIRNGFRSH ESPGEMDLNDDWYLLSKDEVAYLKKEFEKH GAVVNESHEDALVEDIFDKEDEDKDOFFISAR FTYKHIBEL 793 2143 A 6446 3201 152 PRIKRI-VYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEWSKSTHCPSVFLACAWSCRNLLAFTM DLRSDQODLTRMIHILDTEHPWDLHSINSEHH EAITCLEWDQSGFPGFISRWFTGGGNI VATADGSSASPVOPYKVCVSVVSEKCRIDT DLLPSLEMRCTTDLARKOKFPAITHLKFLARD MSEQVILCASSOTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQFTILKWRLISATNDLRVSS VALPKLIPISLTNTDLKVASDTQFYPGGLGLAL AFHDOSVHITHRISLGTMRAVYSSAARPPYD							
DAELLWRLARASRDVAQLSRTSEEEKKLLVY EALEYAKRA/JEKNESSFASHKWA/AICLSDV GDYEGIKAKIANA/TIKEHFEKAIELNPKDATS IIHLMGIWCYTFAEMPWYQRRIA*NACLQLPP **PPPYPEKAIGLYPHRAEQUPNPYSKNILLLIG KTYLKLHNKKLAFWIMKAKDYPAHTEED KQIQTEAQLLTSPSEKN 791 2141 A 6434 3 1460 IALLIVDGLAWDDQGGLALLHISPSKLIL*QDS SGMS/YVMVRCTITRAFKSLLCHICQYSIGPQ **VTCPQQDACKF*KSTAN*GG*RE**PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEGRQQQSGEAELALARYSISINGLSNILN NETSGTYANGSVIDLFKSEGYYNVVSGQPSP DQSGLDMTGIKQJKQEPIYDLTSVPNLFTYNS FNNNQLAPGITNMTEIDRIAQNIKSSILETCQY TMEELHQLAWQTHTYEEIKAYQSKSEALW QQCAIQTHAIQTVVEFAKRTGFMELCQNDQ IILLIKSGCLEVVLVRNMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNBFAKNLCSL QTEEEIALFSSAVLISPDRAWLIEPKVQKLQ EKIYFALQHTIQKNHLDBITLAKLLAKPITTA VCNLHGEKLQVFKQSHPEIVNTPPLYFELT NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALJPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNQQPIWFT LIGIEALKGWGPGA*KDMCVGEKKLIIPPA LGYGKBGKGKUPPSTLIFNIDLLEIRNGFRSH ESGCEMDLNDDWKLSKDEVKJKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR ETTYKHDEL 793 2143 A 6446 3201 152 PRIKRIVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEWMSKSTHOWKLSKDFFWHIGDISCH* GAVVNESHHDALVEDIFDKEDEDKDGFISAR ETTYKHDEL BETTYKHDEL URSDDQDLTRIMHILDTEHPWDLHSIPSHH EATTCLEVDQSGFFGHI-FSWLYSGHVFS SGQVLTSTESLCRLRARVALADIAFTGGMI VVATAGOSSAUSFVOFTKVCVSVVSECKDIDT DLIPSLFMRCTTDLNKKDKFPATHLKFLARD MSEQVLLCASSQTSSIVEGGBFILKVRSI SGQVLTSTESLCRLRARVALADIAFTGGMI VVATAGOSSAUSFVOFTKVCVSVVSECKDIDT DLIPSLFMRCTTDLNKKDKFPATHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGELPNNI FQQISPVVGDKQFILKWRLSLATNDLDRVSS VALPKLIPISLTNTDLKVASDTQFYPGGIGLUS VALPKLIPISLTNTDLKVASDTQFYPGGIGLAL AFHDOSVHIVHRISLSLTMMVYSSAARPPYD	1	1	ł				
BALEYARRAI/JEKNESSÄASHKWYAICLSDV	I	1	1	!		l	
GDYEGIKAKIANAYJIKEHPEKAJELIPKIDATS IHLMGIWCYTFAEMPWYQRRIA*NACLQLPP *PPPPYEKALGYFRAEMPWYQRRIA*NACLQLPP *PPPPYEKALGYFRAEMPWYQRRIA*NACLQLPP *PPPYEKALGYFRAEMPWYQRRIA*NACLQLPP *PPPYEKALGYFRAEMPWYQRRIA*NACLQLPP KTYLKLHNKKI.AAPWLMKAKDYPAHTEED KQIQTEAQLITSPISEM. 791 2141 A 6434 3 1460 IALLIVDGLAWDDQGGLALLHISPSKILL*ODS SGMSYVMYRCITIRAFFKSLLCHICVSIGPO *VTCPQQDACKE*KSTANGG*RE**PPQVLFF AFLSNPAVKFGRNSKKQRDSLVAEVQKHQQ RLQEQRQOQSGAEALANNGG*RE**PPQVLFF AFLSNPAVKFGRNSKKQRDSLVAEVQKHQQ RLQEQRQOQSGAEALANNGGSLSNLN NETSGTYANGSVIDLPKSEGYYNVVSQCPSP DQSGLDMTGIKQKQEPIYDLTSVPNLFTYNSS FNNIGGLARGTITMTEDIRJNIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQTHAIQYVVEFAKRITGFMELCONDQ ILLKKSGCLEVVLVRMCRAFNPINNTVLFEG KYGGMQMKKALGSDDLVMEADFAKNLCSL QLTEEIALRSSAVLISPRAWLEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFFCSNMRJFLWNAVLTLFV LGILEALKGWGGAKYLGKFFCFRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNQPIWFT LGILEALKGWGGAKYLGKFFCFRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNQPIWFT LGILEALKGWGGAKYLGKFFCFRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNQPIWFT LGILEALKGWGGAKYLGKFFCFRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNQPIWFT LGILEALKGWGGAKYLGKFFCFRKTKGGDL MLVHYEGYLEKDGSLFFSTHKHNNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKFFEKH GAVVNESHHDALVEDIFDKEDBCAGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRIKRILVYTEEDGGARPEALGKIAPRTPAELG AKADQELVTALMCDLRSPAAGGGMDLAYV CEWEKWSKSTHCPSVFLACAWSCRNLLAFTM DLRSDDQDLTRMIHLIDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGI-FSRWPTGQKKCWS MGVSTLAINSWESSVGSLWEGGFH,WALS, WLNNGYKLALHVEKSGASFGEKFSRWFS PISITLPGGNAMEGWIAVTVSGLVTVSLLQIP SGQVLVTSTESLCRLRARVALADIAFTGGGN VVATADGSSASSPQFFKVCVSVVSEKCRIDT DILPSLFMRCTTDLINRDKFPATHIKHLARD MSEQVILLCASSGTSSICEWSCERFSRWFS PISITLPGGNAMEGWIAVTVSGLVTVSLLQIP SGQVLVTSTESLCRLRARVALADIAFTGGGN VVATADGSSASSPQFVFKVCVSVVSEKCRIDT DILPSLFMRCTTDLINRDKKPPATHIKHLARD MSEQVYLLCASSGTSSICEWSCERFSRWFS VVATADGSSASSPQFFVKVCVSVVSEKCRIDT DLPSLFMRCTTDLINRDKKPPATHIKHLARD MSEQVYLLCASSGTSSICEWSCERFSRWFS VVATADGSSASSPQFFVKVCVSVVSEKCRIDT DLPSLFMRCTTDLINRDKKPPATHIKHLARD MSEQVYLLCASSGTSSICEWSCERFSRWFS VALPEUSCH	1		İ	1			
IHLMGIWCYTFAEMPWYQRRIA*NACLQLPP *FFPYEKALGVYFHRAEQVDPNFYSKNLLLLG KTYLKLINKKLAAFWLKKAKDYPAHTEED KQIQTEAQLLTSTSEKN A 6434 3 1460	1		l]		l	EALEYAKRA/L/EKNESSFASHKWYAICLSDV
IHLMGIWCYTFAEMPWYQRRIA*NACLQLPP *FFPYEKALGVYFHRAEQVDPNFYSKNLLLLG KTYLKLINKKLAAFWLKKAKDYPAHTEED KQIQTEAQLLTSTSEKN A 6434 3 1460	1		l]		1	GDYEGIKAKIANAYIIKEHFEKAIELNPKDATS
#FPPYEKALGVPHRAEQÜDPNFYSKNLLLIG KTYLKLHNKKLAAFWLMKAKDYPAHTEED KQIQTEAAQLLTSFSKIN 791 2141 A 6434 3 1460 IALLIVDGLAWDQGGLALLHISPSKLIL-YODS SGMS/YVMYRCTITRAFKSLI.CHICQYSIGPQ *VT\CPGQDACKE*KSTAN*GG*RE**PQVLFF AFILSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQGGEALARVYSSSISNGLSNLN NETSGTYANGSVIDLPKSEGYYNVVSQOPSP DQSGLDMTGIKQRJCKJCSUCYNVVSQOPSP DQSGLDMTGIKQRJCKJCSUCYNVVSQOPSP DQSGLDMTGIKQRJCKJCSUCYNVVSQOPSP DQSGLDMTGIKQRJCKJCSUCYNVVSQOPSP DQSGLDMTGIKQRJCKJCSUCYNVVSQOPSP DQSGLDMTGIKQRJCKJCSUCYNVVSQOPSP DQSGLDMTGIKQRJCKJCSUCYNVVSQOPSP DQSGLDMTGIKQRJCKJCSUCYNVVSQOPSP DQSGLDMTGIKQRJCKJCSUCYNVVSQOPSP DQSGLDMTGIKQRJCKJCSUCYNVVSQOPSP DQSGLDMTGIKQRJCKJCSUCYNVVSQOPSP DQSGLDMTGIKQRJCSUCYNVVSQOPSP DQSGLDMTGIKQRJCSUCYNVVSQOPSP DQSGLDMTGIKQRJCSUCYNVVSQOPSP DQSGLDMTGIKQRJCSUCYNVSSSSREALW QQCAIQTHAIQTVVERKTIGFMELCONDO IILLKSGCLEVVLVRMCRAFNIGFMELCONDO IILLKSGCLEVVLVRMCRAFNIGFMELCONDO IILLKSGCLEVVLVRMCRAFNIGFMELCONDO IILLKSGCLEVVLVRMCRAFNIGFMELCONDO IILLKSGCLEVVLVRMCRAFNIGFMELCONDO IILLKSGCLEVVLVRMCRAFNIGFMELCONDO IILLKSGCLEVVLVRMCRAFNIGFMELCONDO IILLKSGCLEVVLVRMCRAFNIGFMELCONDO IILLKSGCLEVVLVRMCRAFNIGFMELCONDO IILLKSGCLEVVLVRMCRAFNIGFMELCONDO IILLKSGCLEVVLVRMCRAFNIGMELCONDO IILLKSGCLEVVLVRMCRAFNITAMELTTVLS INTERCTORDO IILLKSGCLEVVLVRMCRAFNITAMELTTVLS INTERCTORDO IILLKSGCLEVVLVRMCRAFNITA	1	1	!	[1	
791 2141 A 6434 3 1460 IALLIVOGLAWDOGGCALLHISPSKLIL*QDS SGMSYYMVRCTITRAFKSLLCHICQYSIGPQ *VTCPGQDACKE*SKNGG*PS*PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQSGEAEALARYYSSSISNGLSNLN NETSGTYANGSVIDL*PKSEGYYNVVSQOPSP DQSGLDMTQHKQIKQEPPYDLTSYPNLFTYNSS FNNGQLAPGITM*EIDRIAQNIKSHLETCQY TMEELHQLAWQTHITYEHKAYQSKSREALW QQCAIQTTHAIQYVVEFAKRITGFMELCQNDQ ILLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEEIALFSSAVLISPDRAWLEFRKYQKLQ EKIYFALQHVIOKNHLDETLAKLLAKHPITTA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTTRCTCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGHLEALKGWGPGA*KDMCVGEKRKLIHPPA LGYGKEGKGKPPESTLIFNDLLEIRNGPRSH ESPGEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHIDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRIKRITVYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRILLAFTM DLRSDDQDLTRMIHILDFHPWDLHSIPSEHH EAITCLEWDQSGFFGFIFSRWPTGQIKCWS MGVSTLAINSWESSVGSLVEGGPHLWALSI WLHNIGVKLALHVEKSGASSFGEKFSRIVKFS PISLTLFGGNAMEGWIAVTVSGLVTVSLLQIP SGQVLTSTBSLCCIR,ARTAVALADIAFTGGRH VVATADGSSAISPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAHLKFLARD MSGQVLLCASSOTISSTCRARVALADIAFTGGRH VVATADGSSAISPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAHLKFLARD MSGQVLLCASSOTISSTCWSLKREGIEPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VALPKLPISLTNTDLKKVASDTGFYPGLGLAL AFHDGSVHIVHRLSLQTMAYFYSSAAPRPVD	1			İ			
RCIOTEAAQLLTSFSEKN 791 2141 A 6434 3 1460 1ALLIVDGLAWDDQGGLALLHISPSKLIL*QDS SGMSYYMVRCTITRAFFKSLLCHICQYSIGPQ *VTICPGQDACKE*KSTAN*GG*RE**PQVLFF AFLSNPAVKFGMSKKGROSLYAEVQKHQO RLQEQRQQQSGEAEALARVYSSSISNGLSNLN NETSGTYANGSVIDLEKSEGYYNVSQOPPS DQSGLDMTGKIQRQKORDSLYAEVQKHQO RLQEQRQQQSGEAEALARVYSSSISNGLSNLN NETSGTYANGSVIDLEKSEGYYNVSQOPPS DQSGLDMTGKIQRQKORDSLYAEVQKHQO RLQEQRQQQSGEAEALARVYSSSISNGLSNLN NETSGTYANGSVIDLEKSEGYYNVSQOPPS DQSGLDMTGKIQRQKOPSLTSVPNLFTYNSS PNNGQLAPGITNMTEIDRIAQNIKSHLETCQY TMEELHQLAWQTHTYEIKAYQSKSREALW QQCAQITHAIQYVVEFAKRITGFMELCQNDQ ILLIKSGCLEVVLVRMCRAFNPLNTVLFEG KYGGMQMFKALGSDN WAEAFPENNTLFPRNTVLFPLYKELF NPDCATACK VCNLHGERLQVFKQSHLIPPNLNEFPLYKELF NPDCATACK SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEFEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNQQPIWFT LGILEALKGWGPGA*KDMCVGEKRKLIPPA LGYGKEGKGKIPPESTLIFNIDLLERNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHIDALVEDIFDKEDEDKDGFISAR EFTYKHDEL TGLEALKGWGPGA*KDMCVGEKRKLIPPA LGYGKEGKGKIPPESTLIFNIDLLERNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHIDALVEDIFDKEDEDKDGFISAR EFTYKHDEL TGLEALKGWGPGA*KDMCVGEKRKLIPPA LGYGKEGKGKIPPESTLIFNIDLLERNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHIDALVEDIFDKEDEDKDGFISAR EFTYKHDEL TGLEALKGWGPGA*KDMCVGEKRKLIPPA LGYGKEGKGKIPPESTLIFNIDLSHINAT DLRSDDQDLTRMIHLDFHWDLHSTSSHH EAITCLEWDQSGFPGIFSRWFTS PSLTLFGGNAMEGWIAVTVSGGPHLWALS WLHNGWKLALHVERSGASSFGEKFSRIVKFS PSLTLFGGNAMEGWIAVTVSGQFHUWALS WLHNGWKLALHVERSGASSFGEKFSRIVKFS PSLTLFGGNAMEGWIAVTVSGCHLWERSCHUET DILPSLFMRCTTDLNKKDKFPAITHLKFLARD MSEQVLLCASSOTSSUCWSLKEGELDT DILPSLFMRCTTDLNKKDKFPAITHLKFLARD MSEQVLLCASSOTSSUCWSLKEGELDT DILPSLFMRCTTDLNKKDKFPAITHLKFLARD MSEQVLLCASSOTSSUCWSLKEGELDT DILPSLFMRCTTDLNKKDKFPAITHLKFLARD MSEQVLLCASSOTSSUCWSLKEGELDT DILPSLFMRCTTDLNKKDKFPAITHLKFLARD MSEQVLLCASSOTSSUCWSCHEGELDT AFHDGSVHIVHRLSLOJOMAVFYSSAAAPRVVD	1	ł					
791 2141 A 6434 3 1460 IALLIVDGLÄWDDQGGLALLHISPSKLIL-ODS SGMS/YVMVRCTITRAFFKSLLCHICQYSIGPQ *VTICPGQDACK#-KSTAN*GG*RE***PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQSGEAEALARVYSSSISNGLSNLN NETSGTYANGSVUDJKSSEGYYNVVSGQFSP DQSGLDMTGIKQIKQQEIYDUTSVPNLFTYSS FNNGQLAPGITMTEIDRIAQNIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAQITHAIQYVVETAKRTGFMELCQNDQ ILLKSGCLEVVLVMCCRAFIPPLINTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEESIALFSSAVLISPDRAWLIEPRKVQKLQ EXIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFFCFSNMRLFLWNAVLTLFV TSLIGALIPEPSVLEVLQKFFICHRKTKGGDL MLVHYEGYLEKGSLFISTHKHNNOQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAVLKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRIKRI.VVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMHILDTEHPWDLHSIPSEHH EAITCLEWDDSGFFGLIFSTHKHNAVLAVLSV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMHILDTEHPWDLHSIPSEHH EAITCLEWDDSGFFGLIFSREHWALS WHINGVKLALHVESGASSFGEKFSRIVKFS PISLTLFIGGNAMEGWIAVTVSGLVTVSLLQP SGVVLYSTLEGGPHLWALS WHINGVKLALHVESGASSFGEKFSRIVKFS PISLTLFIGGNAMEGWIAVTVSGLVTVSLLQP SGVVLYSTSEGLEGIPVNNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRXDKFPAITHLKFLARD MSGVULCASSOTSSIVECWSCREGIPVNNI FQQISPVGBKQPTILKWRLSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDOSVHTVIRLSLQTMAVFYSSAAPRPVD	1	İ		1			
SGMS/YVMVRCTITRAFFKSLLCHICQYSIĞPQ *VTUCPGQÜACKE*KSTAN*GG*RE**PQVLIFF AFLSNPAVKFÇRKMŞKKQRDSLYASPQKHQQ RLQEQRQQQSGEAEALARVYSSSISNGLSNIN NETSGTYANGSVIDLPKSEGYYNVVSGQPSP DQSGLDMTGIKQIKQEPIYDLTSVPNLFTYSS FPNNGQLAPGITMTEIDRIAQNIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ ILLLKSGCLEVVLVRMCRAFNPINNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKPPITTA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIEPSPVKIEVLQKFPICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*KDMCVGEKRILIPPA LGYGKEĞKĞKIPETJENDILLEIRIGPPSH ESFQEMDLNDDWKLSKDEVKAYLKEFEKH GAVVNESHIDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRLVVTEFDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTTLGSPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGPFGPLFSRWFTGQIKCWS MGYSTLANISWESSVGSLIVEGGPHLWALSI WLHNGVKLALHVEKSGASSFGEKFSRIVKFS PSLTLFGGRAMEGWIAVTVSCLVTVSLLQP SGQVLIATSTESLCRLRARVALADIAFTGGGNI VVATADGSSASPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLINRKDKFPAITHLKFLARD MSEQVLLCASSGTSSIVECWSLRKEGEIPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VVALPKLIPISLTNTILKVASDTQFYPGLGLAL AFHDGSVHIVHILSLQTMAVFYSSAAPRPVD	701	2141	_	6424	2	1460	
#VITCPGQDACKE*KSTAN*GG*RE**PQVLFF AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ RLQEQRQQQSGEALARVYSSSISNGLSNLN NETSGTYANGSVDLPKSEGYYNVVSGQPSP DQSGLDMTGIKQIKQEPIYDLTSVPNLFTYVSS FNNGQLAPGITMTEIDRIAQNIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQTHAIQYVVEFAKRITGFMELCQNDQ IILLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNBAFDFAKNLCSL QLTEEGLALFSSAVLISPDRAWLIEFRVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFGDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKFFICHRKTKGGDL MLVHYEGYLEKGGLFISTHKHNNGQPIWFT LGILEALKGWGPGA*KDMCVGEKRLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHIDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRIKRILVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHGSVPLACAWSCRALLAFTIM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFFGFLFSRWFTGGIKCWS MGYSTLANISWESSVGSLVEGGPHLWALSI WLHNGVKLALHVEKSGASSFGEKFSRIVKFS P\SLTLFGGRAMEGWIAVTVSGLVTVSLQPI SGQVLISTTESLCLARAVALADIAFTGGGI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSGQVLICASSGTSSIVECWSLRKEGELPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA V\ALPKLIPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHELSLQTIMAVFYSSAAPRPVD	'71	2141	^	0434	3	1400	
AFLSNPAVEFGRMSKKQRDSLYAEVQÜHQQ RLQEQRQQQSGAEALARYYSSSISNGLSNIN NETSGTYANGSVIDLPRSEGYYNVVSGQPSP DQSGLDMT\GIKQIKQPPIDLTSVPNLFTY\SS FN\NGQLAPGIT\MTEIDRIAQ\NIK\SHLETCQY T\MEELHQLAWQTHT\YEEIKAYQ\SK\SREALW QQCAIQITHAIQ\YV\FAK\RITGFMELCQ\NDQ ILLLK\SGCLEV\LVR\MCRAF\PPL\N\T\UFF\LEG K\YGG\MQ\MFKALG\SDDL\V\\AE\ALI\ETP\R\V\\C\\LQ\L\ Q\LTA\LAG\SAV\LISP\BRA\WLIEPR\LY\\LQ\L\ EKI\YFALQ\HV\Q\K\\AH\DDETLAKLIA\HT\T\ V\C\\LH\GEKL\Q\V\K\Q\SH\PPL\\T\\LF\\LY\\LY\\LY\\LE\\LY\\LY\\LIP\\LY\\LY\\LY\\LIP\\LY\\LY\\LIP\\LY\\LY\\LIP\\LY\\LY\\LIP\\LY\\LY\\LY\\LIP\\LY\\LY\\LIP\\LY\\LY\\LIP\\LY\\LY\\LIP\\LY\\LY\\LIP\\LY\\LY\\LIP\\LY\\LY\\LIP\\LY\\LY\\LIP\\LY\\LY\\LY\\LY\\LIP\\LY\\LY\\LY\\LIP\\LY\\LY\\LY\\LY\\LY\\LY\\LY\\LY\\LY\\L	1						
RLQEQRQQSGEAŁALARYYSSSISNIGLSNI.N NETSGTYANGSVIDLPKSEGYYNVVSQQPSP DQSGLDMT\GIKQIKQEPIYDLTSVPNLFTYYSS FNNGQLAPGITIMTEIDRIAQNIKSHLETCQY TIMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ ILLLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDJVNEAFDFARNLCSL QLTEEIALFSSAVLISPDRAWLIEPRKYQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK AVCHHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK LGULEALKGWGPGA*KJDMCYGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGQPIWFT LGILEALKGWGPGA*KJDMCYGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAYVNESHIDALVEDIFDKEDEDKDGFISAR EFTYKHBEL 793 2143 A 6446 3201 152 PRIKRI.VYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRPPAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGPFGFLFSRWPTGQIKCWS MGVSTLAINSWESSVGSLIVEGGPHLWALSI WLHNIOGVKLALHVEKSGASFGEKFSRIVKFS PISLTLFIGGNAMEGWIAVTVSGLVTVSLLQP SGQVLYTSTESLCRLRARVALADIAFTGGGNI VVATADGSSAISPVQFVKVCVSVVSVSEKCRIDT DILPSLFMRCTTDLINRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQFTILKWRLSATNDLDRVSA VALPKLPISLTNITDLKVASDTQFYPGLGLAL AFHDGSVHIVHRISLQTMAVYYSSAAPRPVD		1	1				
NETSGTYANGSVIDLPKSEGYYNVVSGQPSP DQSGLDMTGIKQIKQPIYDLTSVPNLFTYSS FNNRQLAPGTTMTEIDRIAQNIIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQITHAIQYYVEFAKRITGFMELCQNDQ ILLLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLLAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFFCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYJLEKDGSLFHSTHKHNNNQPIWFT LGILFALKGWGFGA*KDMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESSQEMDLNDDWKLSKDEVKAYLKEFEKH GAVVNESHIÐALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRIKRI-VYTEEDGGARFEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLLAFTM DLRSDDQDLTRMHILDTEHPWDLHSIPSEHH EAITCLEWDQSGPFGFLFSRWPTGGIKICWS MGVSTLAINSWESSVGSLIVEGGPHLWALSI WLHNGVKLALHVEKSGASSFGEKFSRIVKFS PISLTLFGGNAMEGWIAVTVSGLUP SGQVLITSTESLCRLRARVALADIAFTGGGNI VVATADGSSASPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTIMAVFYSSAAPRPVD	1		l				AFLSNPAVKFGRMSKKQRDSLYAEVQKHQQ
DOSGLDMTGIKQIKQEPIYDLTSVPNLFTYNSS FNNGQLAPGITMTEIDRIAQNIIKSHLETCQY TMEELHQLAWQTHTYEEIKAYQSKSREALW QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ IILLKSGCLEVVLVMACRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVMEAPDFAKNLCSL QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKFFICHKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESSQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRILVYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLLAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIKCWS MGVSTLAINSWESSVGSLVEGGPHLWALS, WLHNGVKLALHVEKSCASSFGEKFSRIVKFS PSSTLFFGGNAMEGWIAVTVSGLVTVSLLQP SGQVL\TSTESLCRLRARVALADIAFTGGGNI VVATADGSSASPVGTYKVCVSVVSEKCRIDT DILPSLFMRCTTDLINKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLKKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVRIKSLQTMAVFYSSAARPPVD	1		١.	l			
FNNGQLAPGITMTEIDRIAQNIKSHLETCQY TIMELHQLAWQTHTYEIKAYQSKSREALW QCQAIQITHAIQYVVEFAKRITGFMELCQNDQ ILLLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*KJDMCVGEKRKLIIPPA LGYGKEGKGKPPESTLIFNDLLEIRNGFRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRIKRLVYTEFDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTMIHILDITEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIKCWS MGVSTLANSWESSVGSLVEGGPHLWALSI WLNNGVKLALHVEKSGASSFGEKFSRIVKFS PSSLTLFIGGNAMEGWIAVTVSGLVTVSLLQP SGQVLITSTESLCRLRARVALADIAFTGGGNI VVATADGSSANSPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLINRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVRIKSLQTMAVFYSSAARRPVD							NETSGTYANGSVIDLPKSEGYYNVVSGQPSP
FNNGQLAPGITMTEIDRIAQNIKSHLETCQY TIMELHQLAWQTHTYEIKAYQSKSREALW QCQAIQITHAIQYVVEFAKRITGFMELCQNDQ ILLLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*KJDMCVGEKRKLIIPPA LGYGKEGKGKPPESTLIFNDLLEIRNGFRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRIKRLVYTEFDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTMIHILDITEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIKCWS MGVSTLANSWESSVGSLVEGGPHLWALSI WLNNGVKLALHVEKSGASSFGEKFSRIVKFS PSSLTLFIGGNAMEGWIAVTVSGLVTVSLLQP SGQVLITSTESLCRLRARVALADIAFTGGGNI VVATADGSSANSPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLINRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVRIKSLQTMAVFYSSAARRPVD	1				•		DOSGLDMT\GIKOIKOEPIYDLTSVPNLFTY\SS
TMEELHQLAWQTHTYEEIKÄYQSKSREALW QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ IILLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEEIALFSSAVLISFDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWBPPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRIKRILVYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHLDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIKICWS MGVSTLAINSWESSVGSLIVEGGPHLWALSI WLHNIGVKLALHVEKSGASSFGEKFSRIVKFS PSSLTLFGGNAMEGWIAVTVSGLVTVSLLQP SGQVLTSTESLCRLRARVALADIAFTGGGNI VVATADGSSASPVCFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFFYPGLGLAL AFHDGSVHVHRLSLQTMAVFYSSAAPRPVD	1	ļ		1			
QQCAIQITHAIQYVVEFAKRITGFMELCQNDQ IILLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKFFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNQPIWFT LGILFALKGWGPGA*KJMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLJAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIKICWS MGVSTLANSWEISSVGSLVEGGPHLWALSI WLHNIGVKLALHVEKSGASSFGEKFSRIVKFS PSLTLFGGNAMEGWIAVTVSGLVTVSLLQIP SGQVLTTSTESLCRLRARVALADIAFTGGGNI VVATADGSSASPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTOFYPGLGLAL AFHDGSVHVHRLSLQTMAVFYSSAAPRPVD			1	1			
ILLKSGCLEVVLVRMCRAFNPLNNTVLFEG KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLTEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPITTA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFFGFLFSRWPTGQIKICWS MGYSTLANSWESSVGSL/VEGGPHLWALS WLHNGVKLALHVEKSGASSFGEKFSR/VKFS PSLTLFNGGNAMEGWIAVTVSGLVTVSLLQP SGQVLTSTESLCRLRARVALADIAFTGGGNI VVATADGSSANSPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD		ļ	l	1			
KYGGMQMFKALGSDDLVNEAFDFAKNLCSL QLITEEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVNNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRILKRI.VVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFILFSRWPTGQIKICWS MGVSTLAINSWEISSVGSLIVEGGPHLWALSI WLHNNGVKLALHVEKSGASSFGEKFSRIVKFS PISLTLFIGGNAMEGWIAVTVSGLVTVSLLQP SGQVLITSTESLCRLARAVALADIAFTGGGNI VVATADGSSAISPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLINKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLISATNDLDRVSA VVALPKLPISLTNTDLKVASDTOFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD		1	l				
QLTEEEIALFSSAVLISPDRAWLIEPRKVQKLQ EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFINDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRLVYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFFSRWPTGQIK\CWS MGYSTLAINSWESSVGSLVEGGPHLWALS\ WLHNGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTV\SGLVTV\SLLQ\P SGQVL\TST\EST\ESLCRLRARVALADIAFTGGGNI VVATADGSSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VVALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD		1	l				
EKIYFALQHVIQKNHLDDETLAKLIAKIPTITA VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTRCFGCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRI.VYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACAWSCRNLIAFIM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIKVCWS MGVSTLAINSWEISSVGSLIVEGGPHLWALS\ WLHINGVKLALHVEKSGASSFGEKFSRIVKFS P\SLTLF\GGNAMEGWIAVTV\SGLVTVSLLQ\P SGQVL\TSTESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1	ľ	l				
VCNLHGEKLQVFKQSHPEIVNTLFPPLYKELF NPDCATACK 792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRI.VVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK/CWS MGVSTLAINSWEISSVGSLIVEGGPHLWALSI WLHINGVKLALHVEKSGASSFGEKFSRIVKFS PISLTLFIGGNAMEGWIAVTVSGLVTVSLLQIP SGQVL/TSTTESLCRLRARVALADIAFTGGGNI VVATADGSSAISPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVILCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VALPKLPISLTNTDLKVASDTOFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1		l				QLIEEEIALFSSAVLISPDRAWLIEPRKVQKLQ
792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHIDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRI.VYTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIKVCWS MGVSTLAINSWESSVGSLVEGGPHLWALSI WLHNGVKLALHVEKSGASSFGEKFSRVKFS PSSLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1	Ī	J	J			
792 2142 A 6440 92 781 SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRI.VVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLAINSWESSVGSLVEGGPHLWALS\ WLHNGVKLALHVEKSGASSFGEKFSRVVKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TSTESLCPLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSGTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD							
TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTILIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC'LEWDQSGFPGFLFSRWPTGQIK'CWS MGVSTLA'NSWE'SSVGSLVEGGPHLWALS\ WLHNGVKLALHVEKSGASSFGEKFSR'VKFS PSLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD		L		<u> </u>			NPDCATACK
TSLIGALIPEPEVKIEVLQKPFICHRKTKGGDL MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*KJMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACAWSCRNLLAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWESSVGSL\VEGGPHLWALS\ WLHNGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	792	2142	A	6440	92	781	SRGTFRCFCRDFFPCFSNMRLFLWNAVLTLFV
MLVHYEGYLEKDGSLFHSTHKHNNGQPIWFT LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRI.VVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLHNIGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD		1					
LGILEALKGWGPGA*K/DMCVGEKRKLIIPPA LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRI.VVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWESSVGSL\VEGGPHLWALS\ WLHNGVKLALHVEKSGASSFGEKFSR\VKFS PSLTLFGGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD]	1					
LGYGKEGKGKIPPESTLIFNIDLLEIRNGPRSH ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD			1]			
ESFQEMDLNDDWKLSKDEVKAYLKKEFEKH GAVVNESHIDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRLVVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TSTESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD		1					
GAVVNESHHDALVEDIFDKEDEDKDGFISAR EFTYKHDEL 793 2143 A 6446 3201 152 PRLKRI.VVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFIM DLRSDDQDLTRMIHILDTEHPWDLHISIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\INSWESSVGSLVEGGPHLWALS\ WLH\INGVKLALHVEKSGASSFGEKFSRIVKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TSTESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1	:	İ	1			
PRIKRI VYTEEDGGARPEALGKIAPRTPAELG ARADQEL VTALMCDL RRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC'LEWDQSGFPGFLFSRWPTGQIK'CWS MGVSTLAINSWESSVGSL'VEGGPHL WALS' WLHINGVKLALHVEKSGASSFGEKFSRIVKFS P'SLTLFGGNAMEGWIAVTVSGLVTVSLLQIP SGQVL/TSTESLCRLRARVALADIAFTGGGNI VVATADGSSAISPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	}		1]			`
793 2143 A 6446 3201 152 PRLKRI.VVTEEDGGARPEALGKIAPRTPAELG ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEK WSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITCLEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLHINGVKLALHVEKSGASSFGEKFSRIVKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1	[l				
ARADQELVTALMCDLRRPAAGGMMDLAYV CEWEKWSKSTHCPSVPLACAWSCRNLIAFIM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD				احتيينا			
CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	793	2143	Α	6446	3201	152	PRLKRLVVTEEDGGARPEALGKIAPRTPÄELG
CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1						ARADQELVTALMCDLRRPAAGGMMDLAYV
DLRSDDQDLTRMIHILDTEHPWDLHSIPSEHH EAITC\LEWDQSGFPGFLFSRWPTGQIK\CWS MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS PSLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD		· '					CEWEKWSKSTHCPSVPLACAWSCRNLIAFTM
EAITC/LEWDQSGFPGFLFSRWPTGQIK/CWS MGVSTLA/NSWE/SSVGSL/VEGGPHLWALS/ WLH/NGVKLALHVEKSGASSFGEKFSR/VKFS P/SLTLF/GGNAMEGWIA/VTVSGLVTVSLLQ/P SGQVL/TST/ESLCRLRARVALADIAFTGGGNI VVATADGSSA/SPVQFYK/CVSVVSEK/CRIDT DILPSLFMRCTTDLNRKDKFPAITHLK/FLARD MSEQVLLCASSQTSSIVECWSLRKEGLP/VNNI FQQISPVVGDKQPTILK/WRILSATNDLDR/VSA V/ALPKLPISLTNTDLK/VASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	ļ	1]				
MGVSTLA\NSWE\SSVGSL\VEGGPHLWALS\ WLH\NGVKLALHVEKSGA\SFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	ł]			į į		
WLH\NGVKLALHVEKSGASSFGEKFSR\VKFS P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRLSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1						
P\SLTLF\GGNAMEGWIAVTVSGLVTVSLLQ\P SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD							
SGQVL\TST\ESLCRLRARVALADIAFTGGGNI VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1		l				
VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA V\ALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD]			
DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1		l]	SGQVL\TST\ESLCRLRARVALADIAFTGGGNI
DILPSLFMRCTTDLNRKDKFPAITHLKFLARD MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1						VVATADGSSA\SPVQFYKVCVSVVSEKCRIDT
MSEQVLLCASSQTSSIVECWSLRKEGLPVNNI FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1						
FQQISPVVGDKQPTILKWRILSATNDLDRVSA VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD		1	1	•			
VALPKLPISLTNTDLKVASDTQFYPGLGLAL AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	1					•	
AFHDGSVHIVHRLSLQTMAVFYSSAAPRPVD	İ		1		•		
	1						
			l				
	L	L	L	L			EPAMKRPKTAGPAVHLKAMQLSWTSLALVG

SEQ ID	SEQ ID	Met	SEQ	Predicted	I Desdies 4	
NO: of	NO: of	hod	ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1 200	in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	1	1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ì	i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
j			1	peptide		/=possible nucleotide deletion, \=possible
			{	sequence		nucleotide insertion
						IDSHGKLSV\LRLSPSMGHPLEVGLALRHLLFL
1		1	1	1	i	LEYCMVTGYDWWDILLHVQPSMVQSLVEKL
	<u> </u>					HEEYTRQTAALQQVLSTRILAMKASLCKLSP
	1	Į				CTVTRVCDYHTKLFLIAISSTLKSLLRPHFLNT
	İ					PDKSPGDRLTEICTKITDVDIDKVMINLKTEEF
1		ł		1	ļ	VLDMNTLQALQQLLQWVGDFVLYLLASLPN
				ŀ		QPCPTSEPCPTSEPSPTSEPSPTSEPSSP*SLC\G
		l		ł		SLLRPGHSFLRDGTSLGMLRELMVVIRIWGLL
İ		1		1	1	KPSCLPVYTATSDTQDSMSLLFRLLTKLWICC
			1			RDEGPASEPDEALVDECCLLPSQLLIPSLDWL
İ						PASDGLVSRLQPKQPLRLQFGRAPTLPGSAAT
ļ		1			ļ	LQLDGLARAPGQPKIDHLRRLHLGACPTEEC
1	ĺ	(ĺ	Í	KACTRCGCVTMLKSPNRTTAVKQWEQRWIK
ļ		[1			NC/LVRWALVAGAPQLPLSPAAPQLLLSYPSA
	1	i			ļ	APEPGCCKSHRSPWTLLGAVNLSPPCRAVEG
			1	1		RGPDACVTSRASEEAPAFVQLGPQSTHHSPRT PRSLDHLHPEDRP
794	2144	A	6490	418	585	NGDKADLENESCRAQVLMPVVPALWEAEGG
			0.50	110	303	GSIEPRDLRLQ*AVITPL\TPAWVTQ
795	2145	A	6499	395	1027	KLLWLPPHSEQKRSPLYHPQGPSGTTPSAP\FS
1			1	"	1027	SHSPPPSLLQA\PSIAAFLRTHGHISASGPLRMP
	ļ					FPH/H*NAFLLVFPGQRSQLTS/PSHYLCREVFP
ĺ		ĺ				DHHHHLCRLSLESSPLFHHRVLFCVPKQNVN
						STRAQIFCLFVHIVGCRCINTFPLHLFRLHLWL
]				•		HFLQIPLCKKNKSVKLGKTVVGRGCQSAAGS
						DTRVRAAVGAPGLPVEPLV
796	2146	A	6503	68	936	HSALLTHSSFCVFTLCQDFFTYSSMSEEVTYA
			1 1			DLQFQNSSEMEKIPEIGKFGEKAPPAPSHVWR
				•		PAALFLTLLCLLLLIGLGVLASMFHVTLKIEM
						KKMNKLQNISEELQRNISLQLMSNMNISNKIR
					i	NLSTTLQTIATKLCRELYSKEQEHKCKPCPRR
	1		1		ł	WIWHKDSCYFLSDDVQTWQESKMACAAQN
						ASLLKINNKNALEFIKSQSRSYDYWLGLSPEE
]					DS/YSWYESG*YNQ\PSAWVIRNAPDLNNMY
	1			[Í	CGYINRLYVQYYHCTYKQRMICEKMANPVQ
797	2147	A	6507	1	881	LGSTYFREA
		**	0307	•	991	PGSTHASARSQVPRSAGEAAPHSRRPPGLLPH
					ĺ	APRAASAQLEERMRDPHPGMTLQEGDCRGS
			į	- 1		QTVSLTMGTADSDEMAPEAPQHTHIDVHIHQ
			{	1		ESALAKLLLTCCSALRPRATQARGSSRLLVAS WVMQIVLGILSAVLGGFFYIRDYTLLVTSGA
	l	l	1	1	1	AIWTGAVAVLAGAAAFIYEKRGGTYWALLR
		1		ļ	l	TLLALAAFSTAIAALKLWNEDFRYGYSYYNS
	•				l	ACRISSSSDWNTPAPTQSPEEVRRLHLCTSFM
		Ί	ſ	1	1	DMLKALFRTLQAMLLGVWILLLASLTPLWL
i		1		ł	ļ	/SL/RGECSQPKG*VPKKRDQKEMLEVSGI*PG
		Į	1	ļ	ļ	STHASARSQVPRSAGEAAPHSRRPPGLLPHAP
				1	[RAASAQLEERMRDPHPGMTLQEGDCRGSOT
1	i		•	Ì	}	VSLTMGTADSDEMAPEAPOHTHIDVHIHOES
	l	1		ł		ALAKLLLTCCSALRPRATQARGSSRLLVASW
l	· .	i	ļ	1	ļ	VMQIVLGILSAVLGGFFYIRDYTLLVTSGAAI
	}	ŀ	ŀ		ŀ	WTGAVAVLAGAAAFIYEKRGGTYWALLRTL
i	į	}				LALAAFSTAIAALKLWNEDFRYGYSYYNSAC
Í	- 1	1	1	- 1	1	RISSSSDWNTPAPTQSPEEVRRLHLCTSFMDM
		1	1			LKALFRTLQAMLLGVWILLLASLTPLWLYC
798	2148		6520	012	2005	WRMFPTKGVSP
.,,,	4140	A	6528	912		VPNYLPSVSSAIGGEVPQRYVWRFCIGLHSAP
ł			[1		RFLVAFAYWNHYLSCTSPCSCYRPLCRLNFG
						LNVVENLALLVLTYVSSSEDF/TWVPG*GRSG

GEO ID	SEQ ID	Met	SEQ	Predicted	Dradioted and	Amino acid sequence (A=Alexino C-Coutoi
SEQ ID NO: of	NO: of	hod	ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A-Alanine C-Cysteine, D-Aspartic Acid, E-Glutamic Acid,
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	l	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	1	ļ		amino acid	of peptide .	T=Threonine, V=Valine, W=Tryptophan,
			i	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
	<u> </u>			sequence		nucleotide insertion
						EVFPEGTGLPLPHSDLPTSWCGHSLQCGSQSS
1						FPPAIHENAFIVFIASSLGHMLLTCILWRLTKK
		1				HTVSQE\DGLSLAGAPRQPRRKSRTSVLRIRV
<u> </u>		1				MVRWELSSNGNPGRGVLGLGLGLGNKLRVV GQNLGL*HCVWVVWETGE*KRWRLQMGIE*
	ļ	1	l			GVASRRQ*VRNSVRGLVCHNSSAPPMYMGFF
			1			SPTVFGGGVGG*LHVTFILHPPEVEAAGIPLLL
						GPSLPQRQGREHIVVILAAPACAPFHDR*WEP
	ł	l				REIRPSP*ELGLRGEPTLSYPASCRVIRQPIP*D
	İ]				RKSYSWKQRLFIINFISFFSALAVYFRHNMYC
j		l				EAGVYTIFAILEYTVVLTNMAFHMTAWWDF
		İ				GNKELLITSQPEEKRF
799	2149	Α	6529	1	874	FFFFQRINFIEHSGSVSLLALACDLGWCEDWS
ĺ	1		1			CCLVQGGGDLVDVVQTNHGEDEAGGDTDSV
			[DEARCKESQQEAQENLREDLCLESFAKDKIL
						QIIEGSEREHEETRTKQAALDGEPLGGGQLTA VHLHPSKEOOGOEGGERORGARTHHWRGW
			}			EKGRRVRLRPPSGKLRADQPVRKLGGPTPS/T
Ì			ŀ			ELPGLQPHAPTPHTA/PATPTYSPAPDTPNPPV
ŀ						RWKCPLPVEPRTRQLCRERTRKACPPKPRPPL
! 		ļ	ł			GLPGDPTGPVTHHAPPVSPTGASGQERRAEP
						GAVSYAHASATK
800	2150	Α	6544	2	662	SAQRWAAVAGRWGCRLLALLLLVPGPGGAS
						EITFELPDNAKQCFYEDIAQGTKCTLEFQVITG
			İ			GHYDVDCRLEDPDGKVLYKEMKKQYDSFTF
			ļ			TASKNGTYKFCFSNE\FSTFTHKTVYFDFQVG
		ļ				E\THLCFLVR/DRVSALTQMESACVSIHEALKS
		}				VIDYQTHFRLREAQGRSRAEDLNTRVAYWSV
801	2151	A	6556	1	1319	GEALILLVVSIGQVFLLKSFFSDKRTTTTRVGS
901	2131	A	0530	1	1319	TPCMECIKGEGLREPQNLSGSQREPQTEGSM DGWRRMPRWGLLLLLWGSCTFGLPTDTTTF
						KRIFLKRMPSIRESLKERGVDMARLGPEWSOP
	İ					MKRLTLGNTTSSVILTNYMDTQYYGEIGIGTP
	ļ	ĺ			;	POTFKVVFDTGSSNVWVPSSKCSRLYTACVY
		١.		1		HKLFDASDSSSYKHNGTELTLRYSTGTVSGFL
		[SQDIITVGGITVTQMFGEVTEMPALPFMLAEF
		ł	i :			DGVVGMGFIEQAIGRVTPIFDNIISQGVLKED
	1					VFSFYYNRDSENSQSLGGQIVLGGSDPQHYE
	1					GNFHYINLIKTGVWQIQMKGVSVGSSTLLCE
		İ				DGCLALVDTGASYISGSTSSIEKLMEALGAKE
	İ					KRLFDYVVKCNEGPTLPPTFLFLLGGKDTPLT
						SADYLFQESYSSKKLSTLAIHAMYIPPPTGPTL VALGATFVIRKFYTEFDRGNNPHGFALAR
802 -	2152	A	6567	13	6147	MCLGRMGASSPRSPEPVGPPAPGLPFCCGGSL
302 .] ^	0507	1.5	V147	LAVVVLLALPVAWGQCNAPEWLPFARPTNL
						TDEFEFFIGTYLNYECRPGYSGRPFSIICLKNS
	1	Ì				VWTGAKDRCRRKSCRNPPDPVNGMVHVIKG
]	Ì		-		IQFGSQIKYSCTKGYRLIGSSSATCIISGDTVIW
	1	l				DNETPICDRIPCGLPPTITNGDFISTNRENFHY
	1					GSVVTYRCNPGSGGRKVFELVGEPSIYCTSND
		l		1		DQVGIWSGPAPQCIIPNKCTPPNVENGILVSD
	1					NRSLFSLNEVVEFRCQPGFVMKGPRRVKCQA
	1	}				LNKWEPELPSCSRVCQPPPDVLHAERTQRDK
						DNFSPGQEVFYSCEPGYDLRGAASMRCTPQG
	1					DWSPAAPTCEVKSCDDFMGQLLNGRVLFPV
						NLQLGAKVDFVCDEGFQLKGSSASYCVLAG
		l .				MESLWNSSVPVCEQIFCPSPPVIPNGRHTGKP LEVFPFGKAVNYTCDPHPDRGTSFDLIGESTIR
	1					CTSDPQGNGVWSSPAPRCGILGHCQAPDHFL
	ŀ				İ	FAKLKTQTNASDFPIGTSLKYECRPEYYGRPF
	L					

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						SITCLDNLVWSSPKDVCKRKSCKTPPDPVNG MVHVITDIQVGSRINYSCTTGHRLIGHSSAECI LSGNAAHWSTKPPICQRIPCGLPPTIANGDFIS TNRENFHYGSVVTYRCNPGSGGRKVFELVGE PSIYCTSNDDQVGIWSGPAPQCIIPNKCTPPNV ENGILVSDNRSLFSLNEVVEFRCQPGFVMKGP RRVKCQALNKWEPELPSCSRVCQPPPDVLHA ERTQRDKDNFSPGQEVFYSCEPGYDLRGAAS MRCTPQGDWSPAAPTCEVKSCDDFMGQLLN GRVLFPVNLQLGAKVDFVCDEGFQLKGSSAS YCVLAGMESLWNSSVPVCEGFCPSPPVIPNG RHTGKPLEVFPFGKAVNYTCDPHPDRGTSFD LIGESTIRCTSDPQGNGVWSSPAPRCGILGHC QAPDHFLFAKLKTQTNASDFPIGTSLKYECRP EYYGRPFSITCLDNLVWSSPKDVCKRKSCKTP PDPVNGMVHVITDIQVGSRINYSCTTGHRLIG HSSAECILSGNTAHWSTKPPICQRIPCGLPPTI ANGDFISTNRENFHYGSVVTYRCNLGSRGRK VFELVGEPSIYCTSNDDQVGIWSGPAPQCIIPN KCTPPNVENGILVSDNRSLFSLNEVVEFRCQP GFVMKGPRRVKCQALNKWEPELPSCSRVCQ PPPEILHGEHTPSHQDNFSPGQEVFYSCEPGY DLRGAASLHCTPQGDWSPEAPRCAVKSCDDF LGGLPHGRVLFPLNLQLGAKVSFVCDEGFRL
				•		KGSSVSHCVLVGMRSLWNNSVPVCEHIFCPN PPAILNGRHTGTPSGDIPYGKEISYTCDPHPDR GMTFNLIGESTIRCTSDPHGNGVWSSPAPRCE LSVRAGHCKTPEOFPFASPTIPINDFEFPVGTS LNYECRPGYFGKMFSISCLENLVWSSVEDNC RRKSCGPPPEPFNGMVHINTDTQFGSTVNYSC NEGFRLIGSPSTTCLVSGNNVTWDKKAPICEII SCEPPPTISNGDFYSNNRTSFHNGTVVTYQCH TGPDGEQLFELVGERSIYCTSKDDQVGVWSS PPPRCISTNKCTAPEVENAIRVPGNRSFFSLTEI IRFRCQPGFVMVGSHTVQCQTNGRWGPKLPH CSRVCQPPPEILHGEHTLSHQDNFSPGQEVFY SCEPSYDLRGAASLHCTPQGDWSPEAPRCTV KSCDDFLGQLPHGRVLLPLNLQLGAKVSFVC DEGFRLKGRSASHCVLAGMKALWNSSVPVC EQIFCPNPPAILNGRHTGTPLGDIPYGKEVSYT CDPHPDRGMTFNLIGESTIRRTSEPHGNGVWS SPAPRCELPVGAACPHPPKIQNGHYIGGHVSL YLPGMTISYTCDPGYLLVGKGFIFCTDQGIWS QLDHYCKEVNCSFPLFMNGISKELEMKKVYH YGDYVTLKCEDGYTLEGSPWSQCQADDRWD PPLAKCTSRTHDALIVGTLSGTIFFILLIFLSWI ILKHRKGNNAHENPKEVAIHLHSQGGSSVHP RTLQTNEENSRVLP
803	2153	A	6574	2	3233	HGRSARLAAVPAEAMPGPRRPAGSRLRLLLL LLLPPLLLLRG\SHAGNLTVAVVLPLANTSY PWSWA\RVGPAVELALAQVKARPDLLPGWT VRTVLGSSENALGVCSDTAAPLAAVDLKWE HNPAVFLGPGCVYAAAPVGRFTAHWRVPLL TAGAPALGFGVKDEYALTTRAGPSYAKLGDF VAALHRRLGWERQALMLYAYRPGDEEHCFF LVEGLFMRVRDRLNITVDHLEFAEDDLSHYT RLLRTMPRKGRVIYICSSPDAFRTLMLLALEA GLCGEDYVFFHLDIFGQSLQGGQGPAPRRPW ERGDGQDVSARQAFQAAKIITYKDPDNPEYL EFLKQLKHLAYEQFNFTMEDGLVNTIPASFH

SEQ ID	SEQ ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uciicc		1	714	amino acid	of peptide	T-Throning Valling Water
1		ł		residue of	sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
i	1		1		sequence	i=1 yrosine, X=Unknown, *=Stop codon,
	Į			peptide		/=possible nucleotide deletion, \=possible
<u></u>	 -	├		sequence		nucleotide insertion
1		İ	ľ		1	DGLLLYIQAVTETLAHGGTVTDGENITQRMW
	1				•	NRSFQGVTGYLKIDSSGDRETDFSLWDMDPE
1					Ì	NGAFRVVLNYNGTSQELVAVSGRKLNWPLG
1		1				YPPPDIPKCGFDNEDPACNQDHLSTLEVLALV
ı						GSLSLLGILIVSFFTYRKMQLEKELASELWRVR
	!	1			'	WEDVEPSSI.ERHI.RSAGSRLTLSGRGSNYGSL
	ľ	1			,	LTTEGQFQVFAKTAYYKGNLVAVKRVNRKR
1						IELTRKVLFELKHMRDVQNEHLTRFVGACTD
ł						PPNICILTEYCPRGSLQDILENESITLDWMFRY
						SLTNDIVKGMLFLHNGAICSHGNLKSSNCVV
		ì				DGRFVLKITDYGLESFRDLDPEQGHTVYAKK
1		}				LWTAPELLRMASPPVRGSQAGDVYSFGIILQE
						IALRSGVFHVEGLDLSPKEIIERVTRGEQPPFR
1						PSLALQSHLEELGLLMQRCWAEDPQERPPFQ
] .			i J			QIRLTLRKFNRENSSNILDNLLSRMEQYANNL
						EEL VEEDTO A VI FEYDY AF ALL YOU BURLEN
						EELVEERTQAYLEEKRKAEALLYQILPHSVAE
}						QLKRGETVQAEAFDSVTIYFSDIVGFTALSAE
1						STPMQVVTLLNDLYTCFDAVIDNFDVYKVET
						IGDAYMVVSGLPVRNGRLHACEVARMALAL
1						LDAVRSFRIRHRPQEQLRLRIGIHTGPVCAGV
1 1			' I			VGLKMPRYCLFGDTVNTASRMESNGEAL\KI
1 1			,			HLSS\ETKAVL\EEFGGFELELRGDVEMKGKG
			122			KYRTYWLLGERGSSTRG
804	2154	Α	6585	2	3837	DAPGRPPVRLPTMELEDGVVYQEEPGGSGAV
[i		1			MSERVSGLAGSIYREFERLIVRYDEEVVKELIP
						LVVAVLENLDSVFAQDQEHQVELELLRDDNE
1 1			1	<i>'</i>		QLITQYEREKALRKHAEEKFIEFEDSQEQEKK
	ì					DLQTRVESLESQTRQLELKAKNYADQISILEE
]			-	'		REAELKKEYNALHQRHTEMIHNYMEHLERT
)	į		1	ļ		KLHQLSGSDQLESTAHSRIRKERPISLGIFPLP
	1		į			AGDGLLTPDAQKGGETPGSEQWKFQELSQPR
l i						SHTSLKDELSDVSQGGSKATTPASTANSDVA
}	1	;		1		TIPTDTPLKEENEGFVKVTDAPNKSEISKHIEV
1						QVAQETRNVSTGSAENEEKSEVQAIIESTPEL
1 1		- 1	1	}		DMDKDLSGYKGSSTPTKGIENKAFDRNTESL
!	ŀ	ŀ	1	1		FEELSSAGSGLIGDVDEGADLLGMGREVENLI
		ŀ	i	1		LENTQLLETKNALNIVKNDLIAKVDELTCEK
	ł	ł	- 1	- 1	ļ	DVLQGELEAVKQAKLKLEEKNRELEEELRKA
		ļ	1		ļ	RAEAEDARQKAKDDDDSDIPTAQRKRFTRVE
	i	l	1			MARVLMERNQYKERLMELQEAVRWTEMIR
1 1	1	j	1	}	· ·	ASRENPAMQEKKRSSIWQFFSRLFSSSSNTTK
	1			ł	1	KPEPPVNLKYNAPTSHVTPSVKKRSSTLSOLP
				1	1	
1 1	J		1.	.]	j	GDKSKAFDFLSEETEASLASRREQKREQYRQ
				- [[VKAHVQKEDGRVQAFGWSLPQKYKQVTNG
	1	- 1	- 1	ļ	l	QGENKMKNLPVPVYLRPLDEKDTSMKLWCA
		- 1	1	1		VGVNLSGGKTRDGGSVVGASVFYKDVAGLD
[[1		i	ľ	1	TEGSKQRSASQSSLDKLDQELKEQQKELKNQ
	İ				- 1	EELSSLVWICTSTHSATKVLIIDAVQPGNILDS
	ľ	1		1	1	FTVCNSHVLCIASVPGARETDYPAGEDLSESG
(1	- 1	. [l	QVDKASLCGSMTSNSSAETDSLLGGITVVGC
	j	ł	ĺ		l	SAEGVTGAATSPSTNGASPVMDKPPEMEAEN
	1	1	1	1	į	SEVDENVPTAEE\ATEATEGNAGSAEDTV\DIS
		- 1	- 1	ļ	ļ	QTGVYTEHVFTDPLG\VQIPEDLSPVYQSSND
	ł			ļ	ľ	SDAYKDQISVLPNEQDLVREEAQKMSSLLPT
					l	MWLGAQNGCLYVHSSVAQWRKCLHSIKLKD
	ļ	}	- 1	J		SILSIVHVKGIVLVALADGTLAIFHRGVDGOW
	ł	1	- 1		l	DLSNYHLLDLGRPHHSIRCMTVVHDKVWCG
	1	ł		ł	}	YRNKIYVVQPKAMKIEKSFDAHPRKESOVRO
				İ		LAWVGDGVWVSIRLDSTLRLYHAHTYQHLO
l i			1		1	DVDIEPYVSKMLGTGKLGFSFVRITALMVSC
						DATE I ASMAITOLOUTOLOU AULUTINIASC

CCC TO	CEOTO	11404	050	D-31-4-3	I 80 Post 1 P	74 H : 0.0
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in NO:	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		,		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
		ì	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	ł	ļ	1	peptide	1	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
	i					NRLWVGTGNGVIISIPLTETVILHQGRLLGLR
	İ	1	ł	1	ļ	ANKTSGVPGNRPGSVIRVYGDENSDKVTPGT
					ĺ	FIPYCSMAHAQLCFHGHRDAVKFFVAVPGQV
						ISPQSSSSGTDLTGDKGRGHLHRSLVVRRP
805	2155	A	6605	469	2602	FGRLLWGTAFKSWKMKAPIPHLILLYATFTQ
	ľ	1				SLKVVTKRGSADGCTDWSIDIKKYQVLVGEP
			1			VRIKCALFYGYIRTNYSLAQSAGLSLMWYKS
	İ	j				SGPGDFEEPIAFDGSRMSKEEDSIWFRPTLLQ
İ						DSGLYACVIRNSTYCMKVSISLTVGENITGL
j						CYNSKMKYFEKAELSKSKEISCRDIEDFLLPT
1						REPEILWYKECRTKTWRPSIVFKRDTLLIREV
ļ						REDDIGNYTCELKYGGFVVRRTTELTVTAPL TDKPPKLLYPMESKLTIQETQLGDSANLTCRA
1	1					FFGYSGDVSPLIYWMKGEKFIEDLDENRVWE
						SDINKILKEHLGEQEVSISLIVDSVEEGDLGNYS
		1				CYVENGNGRRHASVLLHKRELMYTVELAGG
						LGAILLLLVCLVTIYKCYKIEIMLFYRNHFGA
		ļ	i			EELDGDNKDYDAYLSYTKVDPDQWNQETGE
						EERFALEILPDMLEKHYGYKLFIPDRDLIPTGT
]						YIEDVARCVDQSKRLIIVMTPNYVVRRGWSIF
						ELETRLRNMLVTGEIKVILIECSELRGIMNYQE
ļ						VEALKHTIKLLTVIKWHGPKCNKLNSKFWKR
1	ļ					LQYEMPFKRIEPITHEQALDVSEQGPFGELQT
						VSAISMAAATSTALATAHPDLRSTFHNTYHS
ļ	ł	•				QMRQKHYYRSYEYDVPPTGTLPLTSIGNQHT
			ļ			YCNIPMTLINGQRPQTKSSREQNPDEAHTNSA
806	2156	Ā	6614		1504	ILPLLPRETSISSVIW
800	2130	A	0014	3	1584	NSARGGVGVRGARAMATVQEKAAALNLSAL
ļ]					HSPAHRPPGFSVAQKPFGATYVWSSIINTLQT QVEVKKRRHRLKRHNDCFVGSEAVDVIFSHL
i			ì			
	l •					IQNKYFGDVDIPRAKVVRVCQALMDYKVFE AVPTKVFGKDKKPTFEDSSCSLYRFTTIPNOD
						SQLGKENKLYSPARYADALFKSSDIRSASLED
						LWENLSLKPANSPHVNISTTLSPQVINEVWOE
						ETIGRLLQLVDLPLLDSLLKQQEAVPKIPQPK
I	ĺ					RQSTMVNSSNYLDRGILKAYSDSQEDEWLSA
			ĺ	l		AIDCLEYLPDQMVVEISRSFPEQPDRTDLVKE
1						LLFDAIGRYYSSREPLLNHLSDVHNGIAELLV
						NGKTEIALEATQLLLKLLDFQNREEFRRLLYF
1						MAVAANPSEFKLQKESDNRMVVKRIFSKAIV
İ				İ		DNKNLSKGKTDLLVLFL\MDHQKDVFKIPGT
1	[l	ľ	L\HKIVS\VK\LMAIQNGRDPNRDAGYTYCQRI
1				İ	l	DQRDYSNITEKTTIDELLYLLKTLDEDSKLSA
907	0155	_		4100		KEKKK\LLGQFYKCHPDIFIEIIFGD
807	2157	A	6615	4198	2094	FGIVGTFALETDELDSDRDPAIFSLCDFGAMR
]				ļ	ŀ	PQILLLALLTI.GLAAQHQDKVPCKM/VKML
			-		[CPDRVDKKVSCQVLGLLQVPSVLPPDTETLD
]					Í	LSGNQLRSILASPLGFYTALRHLDLSTNEISFL
						QPGAFQALTHLEILSLAHNRLAMATALSAG
						GLGPLPRVTSLDLSGNSLYSGLLERLLGEAPS
]				İ	ļ	LHTLSLAENSLTRLTRHTFRDMPALEQLDLHS
1				ſ		NVLMDIEDGAFEGLPRLTHLNLSRNSLTCISD
						FSLQQLRVLDLSCNSIEAFQTAS\QPQAEFQLT
	ļ. l					WLDLRENKLLHFPDLAALPRLIYLNLSNNLIR LPTGPPGDSKGINAPSEGWSALDLSVAPSGNAS
				l		LPTGPPQDSKGIHAPSEGWSALPLS\APSGNAS
				j	J	GRPLSQLLNLDLSYNEIELIPDSFLEHLTSLCFL NLSRNCLRTFEARRLGSLPCLMLLDLSHNALE
						TLELGARALG\SLRTLLLQGNALRDLPPYTFA
]						NLASLQRLNLQGNRVSPCGGPDEPGP\SGCV\
						AFSGITSLRSLSLVDNEIELLRAGAFLHTPLTE
	LL		اا			A COLLODING DOLADITORECTORA CALCULATION

SEQ ID SEQ ID No: of NO: of NO: of nucl- eotide seq- Note of No: of No: of nucl- eotide seq- Note of No: of No: of nucl- eotide seq- Note of No: of No: of nucl- in nucleotide location corresponding I-Isoleucine, K-Lysine, L- Predicted end nucleotide nucleotide location corresponding I-Isoleucine, K-Lysine, L- Predicted end nucleotide nucleotide location corresponding I-Isoleucine, K-Lysine, L-	
nucl- peptide in nucleotide location F=Phenylalanine, G=Glycin	
	Leucine.
seq- uence 09/496 correspondi to last amino M=Methionine, N=Asparag	zine, P=Proline.
uence 914 ng to first acid residue Q=Glutamine, R=Arginine,	S=Serine,
amino acid of peptide T=Threonine, V=Valine, W	i=Tryptophan,
residue of sequence Y=Tyrosine, X=Unknown, peptide /=possible nucleotide deleti	*=Stop codon,
peptide /=possible nucleotide deletic	on, \=possible
LDLSSNPGLEVATGALG	CI FASI EVI ALOCAL
GLMVLQVDLPCFICLKR	I NI AFNRI CHI PAW
TQAVSLEVLDLRNNSFSI	LLPGSAMGGLETSLR
RLYLQGNPLSCCGNGWI	LAAQLHOGRVDVDA
TQDLICRFSSQEEVSLSH	VRPEDCEKGGLKNI
NLIILTFILVSAILLTTLA	ACCCVRRQKFNQQ
808 2158 A 6619 153 1852 FKALSOVIVINTHI FREA	
TRAESQUITIVIALERED	AAFEVAILLRRMEEG
ARHRNNTEKKHPGGGES	SDASPEAGSGGGGV
ALKKEIGLVSACGIIVGN AGSVGLALIVWIVTGFIT	MOSCIL ASLKO AFEN
PKSGGDYFYVKDIFGGL	ACEL BI MIANI MAN
TNQAVIALTFSNYVLQPI	LFPTCFPPESGLR1 LA
AICLLLLTWVNCSSVRW	ATRVQDIFTAGKLL
ALALIIMGIVQICKGEYF	WLEPKNAFENFOEP
DIGLVALAFLQGSFAYGO	GWNFLNY\VTEELV
DP\YKNL\PRAIFISIP\LVT	
AMSPQELLAS\NAVAVT	
PISVALSTFGGVNGSLFT:	SSRLFFAGAREGHLP
SVLAMIHVKRCTPIPALL MYTLINYVGFINYLFYGV	PTUSTLLMLVISD
DIPRPIKINLLFPIIYLLFW	
CGIGLAIMLTGVPVYFLG	
ELLTLVSQKMCVVVYPE	VERGSGTEEANED
MEEQQQPMYQPTPTKDK	CDVAGQPQP
809 2159 A 6621 1041 223 QDSRKMLPSTSVNSLVQQ	GNGVLNSRDAARH
TAGAKRYKYLRRLFRFR	QMDFEFAAWQMLY
LFTSPQRVYRNFHYRKQ	
VLLSIWLCVSTIGFGFVLI	
VLIDCVGVGLLIATLMWI YDVEWGYAFDVHLNAF	CISING I LVKKŲSKŲ
HVILTDTFIGYLVGNTLW	
GYSVGLLFFS\ALPFLKNT	
LSLALGWNFTHTLCSFYK	
810 2160 A 6623 160 822 SPASGHCRLNGAAVAMF	
QQVAEDKFVFDLPDYESI	
EGMGGSVYFSYPDSNGM	PVWQLLGFVTNGK
PSAIFKISGLKSGEGSQHP	FGAMNIVRTPSVAQ
IGISVELLDSMAQQTPVG	• · · · · · · · · · · · · · ·
QKMLDNFYNFASSFAVS(IPANVVLKWYENFQRRTS	CLEDS! I EMILYMENT
F	O LEL OCCENITA IVIIA
811 2161 A 6627 18 3367 LEGSLNTERAKYYLTITM	PHFTVTKVEDPERG
AAASISQEPSLADIKARIQ	DSDEPDLSONSITG
EHSQLLDDGHKKARNAY	LNNSNYEEGDEYF
DKNLALFEEEMDTRPKVS	SSLLNRMANYTNLT
QGAKEHEEAENITEGKKK	
VYLPCLQNIFGVILFLRLT	
AIVLICCCCTMLTAISMSA	
FMISRALGPEFGGAVGLC	
GAIEIFLVYIVPRAAIFHSD MRVYGTAFLVLMVLVVF	
ACVIVSILAIYAGAIKSSFA	
LSSRHIDVCSKTKEINNMT	
QFFNATCDEYFVHNNVTS	
LWSNYLPKGEIIEKPSAKS	
VDITTSFTLLVGIFFPSVTG	BIMAGSNRSGDLKD
AQKSIPIGTILAILTTSFVY	
VLRDKFGDAVKGNLVVG	
FFSTCGAGLQSLTGAPRLI	LQAIAKDNIIPFLRV

SEQ ID NO: of nucl- eotide scq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion FGHSKANGEPTWALLLTAAIAELGILIASLDL
						VAPILSMFFLMCYLFVNLACALQTLLRTPNW RPRFRYYHWALSFMGMSICLALMFISSWYYA IVAMVIAGMIYKYEYQGAEKEWGDGIRGLS LSAARFALLRLEEGPPHTKNWRPQLLVLLKL DEDLHVKHPRLLTFASQLKAGKGLTIVGSVIV GNFLENYGEALAAEQTIKHLMEAFKVKGFCQ LVVAAKLREGISHLIQSCGLGGMKHNTVVM GWPNGWRQSEDARAWKTFIGTVRVTTAAHL ALLVAKNISFFPSNVEQFSEGNIDVWWIVHDG GMLMLLPFLLK\QHKVWRKCSIRFF\TVAQLE DNSIQMKKDLATFLYHLRIEAEVEVVEMHDS DISAYTYERTLMMEQRSQMLRHMRLSKTER DREAQLVKDRNSMLRLTSIGSDEDEETETYQ EKVHMTWTKDKYMASRGQKAKSMEGFQDL LNMRPDQSNVRRMHTAVKLNEVIVNKSHEA KLVLLNMPGPPRNPEGDENYMEFLEVLTEGL ERVLLVRGGGSEVITIYS
812	2162	A	6628	66	640	AVCTMSEMAELSELYEESSDLQMDVMPGEG DLPQMEVGSGSRELSLRPSRSGAQQLEEEGP MEEEEAQPMAAPEGKRSLANGPNAGEQPGQ VAGADFESEDEGEEFDDWEDDYDYPEEEQLS GAGYRVSAALEEADKMFLRTREPALDGGFQ MHYEKTPFDQLAFIEELFSLMVVNRLTEELG CDEIIDRE
813	2163	A	6630	708	1355	AKMGAYKYIQELWRKKQSDVMRFLLRVRC WQYRQLSALHRAPRPTRPDKARRLGYKAKQ GY/YYIYIGFVFAVIYRIRVRRGGRKRPVPKG ATYGKPVHHGVNQLKFARSLQSVAEERAGR HCGALRVLNSYWVGEDSTYKFFEVILIDPFHK AIRRNPDTQWITKPVHKHREMRGLTSAGRKS RGLGKGHKFHHTIGGSRRAAWRRRNTLQLH RYR
814	2164	Ā	6635	201	1705	KGTEMNKSRWQSRRRHGRRSHQQNPWFRLR DSEDRSDSRAAQPAHDSGHGDDESPSTSSGT AGTSSVPELPGFYFDPEKKRYFRLLPGHNNCN PLTKESIRQKEMESKRLRLLQEEDRRKKIARM GFNASSMLRKSQLGFLNVTNYCHLAHELRLS CMERKKVQIRSMDPSALASDRFNLILADTNS DRLFTVNDVTVGGSKYGIINLQSLKTPTLKVF MHENLYFTNRKVNSVCWASLNHLDSHILLC LMGLAETPGCATLLPASLFVNSHPAGIDRPG\ MLCSFRIPGAWSCAWSLNIQANNCFSTGLSR RVLLTNVVTGHRQSFGTNSDVLAQQFALMA PLLFNGCRSGEIFAIDLRCGNQGKGWKATRLF HDSAVTSVRILQDEQYLMASDMAGKIKLWD LRTTKCVRQYEGHVNEYAYLPLHVHEEEGIL VAVGQDCYTRIWSLHDARLLRTIPSPYPASKA DIPSVAFSSRLGGSRGAPGLLMAVGQDLYCY SYS
815	2165	A	6643	659	3282	NKNILEVPSARTTRIMGDHLDLLLGVVLMAG PVFGIPSCSFDGRIAFYRFCNLTQVPQVLNTTE RLLLSFNYIRTVTASSFPFLEQLQLLELGSQYT PLTIDKEAFRNLPNLRILDLGSSKIYFLHPDAF QGLFHLFELRLYFCGLSDAVLKDGYFRNLKA LTRLDLSKNQIRSLYLHPSFGKLNSLKSIDFSS NQIFLVCEHELEPLQGKTLSFFSLAANSLYSR VSVDWGKCMNPFRNMVLEILDVSGNGWTV DITGNFSNAISKSQAFSLILAHHIMGAGFGFHN IKDPDQNTFAGLARSSVRHLDLSHGFVFSLNS

NO: of No: of peptide coulded uses of the period of the pe	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
SOSAPO SOCION SOSAPO SOCION ient Socient			hod			nucleotide	D=Aspartic Acid, E=Glutamic Acid,
Socion S	1			1		location	
Bell	eotide						
amino acid residue of peptide sequence p		uence		1			M=Methionine, N=Asparagine, P=Proline,
residue of peptide sequence Y=Tyrosine, X=Unknown, **Siop codon, *possible nuclostide declino, *possible nuclostide declino, *possible nuclostide declino, *possible nuclostide declino, *possible nuclostide declino, *possible nuclostide declino, *possible nuclostide declino, *possible nuclostide declino, *possible nuclostide declino, *possible nuclostide declino, *possible nuclostide declino, *possible nuclostide declino, *possible nuclostide, *possible,	uence			914			Q=Glutamine, R=Arginine, S=Serine,
peptide squence possible nucleotide deletion, "possible nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion nucleotide insertion N. QVI.NIA.YNI.LIGEL.YSSENYGI PREVAYIDLO N. QVI.NIA.YNI.LIGEL.YSSENYGI PREVAYIDLO N. QVI.NIA.YNI.LIGEL.YSSENYGI PREVAYIDLO QNAHIALIQOOTERELEEL.QUT.LIQ.ROMALI.THIR FIPISIPIDIE-ISGNICULT.PIPICIL.YNIA.I.PROFESSION TYSENYBE.BCQLF.I.GENMALI.QLAWETSLC.WDV EGL.SHI.C.VI.YLAWPH.QLI.NI.QNAFESS.CSGDQ TYSENYBE.BCQLF.I.GENMALI.YLAWITSLC.PROFESSION N. QVI.NIA.PROFESSION.PIPICIA. ANPOYEVSLS.V.DLITTIRARTICE.CELS.TRIN.WL. N. HITAVITA.QPPA.DIT.CVYPDSILS.OVSL.FIS.HILT.Y GCDEEPLV.RSL.KS.FIS.PLC.YVT.I.TI.L.PHILT.Y TKARGECETE.YX.ARGRU.YR.CHE.QOTE.PROFESSION.PIPIC.PROFESSION.PIPIC.PIPIC.PROFESSION.PIPIC.PIPIC.PROFESSION.PIPIC.PIPIC.PIPIC.PROFESSION.PIPIC.PI	1	1	İ	1			T=Threonine, V=Valine, W=Tryptophan,
Sequence nucleotide insertion				1	1	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
RVPETLEDLKVINLAYNKINELJABEAFYGID N.QVINLSYNLIGEL SYSSPYGIPKYAYTDID (OKNHAIODOTFKTLEKLQITLDIRDNALTTIF FPIPPIDELSGINCLUTLPIKNILTANLIHLSENR LENLDILYFLRVPHLQILTINQNRFSSCSGDO TPSRNSLEQLFLQENMQLAWSTELCWDVA EGLSHLQVILTINHYNLSPROYSBALDISH GISINSNRLTVLSRNDLPANLELDISRNQLL APPIDYPSUSJUDITHNIKPICECHSTRUTH (INTERPEDIT OF SUSSIENTIT GODEEVILSKIKSTEINTRICECHSTRUTH) APPIDYPSUSJUDITHNIKPICECHSTRUTH (INTERPEDIT OKTORITY OKTORITY KYPATLCFSERDFWQSINSLISHIST GODEEVILSKIKSTEINTRICECHSTRUTH) KYPATLCFSSKDFTWQNALKHLDINGNUL KYPATLCFSSKDFTWQNALKTPATLON KYPATLCFSSKDFTWQNALKTPATLON KYPATLCFSSKDFTWQNALKTPATLON KYPATLCFSSKDFTWQNALKTPATLON KYPATLCFSSKQFTWQNALKTP	1	1				}	
MLQVLNLSYNLLGELYSSNYGLPKVAYIDL WHINIOQOTEPELKUQTULDRONALTII FIPSIPDIFLSGNKLVTLPKINLTANLIHLSENR LENLDILVFLLRVPHLQBLINONARSSCSGNO TPSENPSLEGLELGEBNALQLAWETELCWDVW EGLSHRQVLYLNENYLDGLAWETELCWDVW EGLSHRQVLYLNENYLDGLAWETELCWDVW EGLSHRQVLYLNENYLDGLAWETELCWDVW EGLSHRQVLYLNENYLDGLAWETELCWDVW EGLSHRSVLTVLSHDLPANELBLISSNQLL APPPDYFYSLSVLDITHRIKFICECELSTFINWL MHINVTVISLSYLDITHRIKFICECELSTFINWL MHINVTVISLSYLDITHRIKFICECELSTFINWL MHINVTVISLSYLDYCYPDSLSVSSLSSISTE GCDEEEVLKSLKFSLFIVCTVTLTLFLIMTLLTV TKFRGFCICYSTAGRUVKDHPGOTEPDMY KYDAYLCFSSKDFTWVQNALLKHLDTQYSD QNEPRLCFERDPYCGENRANALDDAWINSR KIVCLVSRHFLRDGWCLEAFSYAQGRCLSDL NSALIMVVOUSLSQVOLMKHCSRGFVQKQQ YLRWFEDLQDWGWELHKLSQQLLKKEKEK KDNNIPLQTVATIS BI6 2166 A 6646 I 3811 RDRAGVRPAGKQHAAAAFYDVGGDEPWDS GGARLASLFGLDQAAAGHGNEFPQVTARKQ KRQQTAATGNQATPKTAPATMSTPTLLVAT AVMAYRYNTQGVYKQKGKGAAVGHTTR EYRILLYISQQQPVTVARHHVNELMVRPNNY STFYDDQRQNWSIMERSESKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADQPAVEVGDSLE VAYTGWIPNNYLQQVFDSTANKDKLLRLK LGSGKVIKGWEDMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEYRRVLIA KDSGSDGHSVSRDSAAPSPIPGADNLSADPV VSPTTSIFKSGPFALRTKSNLSSEQLANTSPD AVKAKLISRNAKMGQPMJPLPPQLDSNDSEL EDVNTLQGGGQPVYTSQVSLQANTSPD AVKAKLISRNAKMGQPMJPLPPQLDSNDSEL EDVNTLQGGGGPVYTSQVSLQANTSPD AVKAKLISRNAKMGQPMJPLPPQLDSNDSEL EDVNTLQGGGGPVYTSQVSLQANTSPD AVKAKLISRNAKMGQPMJPLPPQLDSNDSEL EDVNTLQGGGGPVYTSQVSLQANTSPD AVKAKLISRNAKMGQPMJPLPPQLDSNDSEL EDVNTLQGGGGPVYTSQVSLQANTSPD AVKAKLISRNAKMGQPMLPPRQLDSNDSLSL EDVNTLQGGGGPVYTSQVSLQANTSPD AVKAKLISRNAKMGQPMLPPQLDSNDSLSL EDVNTLQGGGGPVTSQVSLQANTSPD AVKAKLISRNAKMGQPMLPPQLDSNDSSL EQUALATIONALSKYLDEFELTDL RESKRTDLLREKKSNAMSCQVSQVYSLAS QCAQAGYSKYSEQCHALASAKDHQKEFTEL QMQLTESKRTDLLREKKSNAMSCQVSQVAAAS QCAQAGYGKIVGULLKKTRVSTDQAAABOLS LVQAELGTARRONGCELLGAKNOMGVSVAAAS DPSELVKKMNQVPGSLRREELEESYNGRT LGTIMTIKMYTLQLAQAGECKELSSSEEEE EKAEERPRPSQQSASASSOQPQAPLARREP ESPMYPSEQVVEBAVPLPQALTTSQDGHR KGDSSPALSEENGEGLEDESTPNLGTT LGTIMTIKMYTLQLAAQAGECELDEI RKSYQEELDKLRQLLKKNDNYGSVPAAAS DPSELVKKMNQVPGSLREELEESYNGRT LGTIMTIKMYTLGLAAQAGAARARPHSQRSSLGDDE SERVAYPSEQVVEBAVPLPQALTTSQDGHR KGDSSAFALSELGGED ED		ļ	ļ		sequence		
QKNHIAIIODQTFKFLEKIQTLDLRDNALTITI FIRSIPBIESONKLYLIKPKINTANLHJSENR LENLDILYFLIKVYHRUITANLHJSENR LENLDILYFLIKVYHLQILINONRFSSCSGDO TYSENSLEQLELGENMQLAWETELGWDVY EGLSHLQVLYLNINYLNSLPPGYFSHLTALR GLSINSNELTVLSHDDLPANLELDISRNQLL APPPDYFVSLSVLDTITINKFICEGELSTFINWL MHTMYTIAGPPADIYCYPDSLSG VSLFSLSTI GCDEEVILKSLKFSLFIVCTVTLTLTAMTILTV TKRRGFCFCYKTAQRLVFKDBPQGTEPDMY KYDAYLCSSKDFTWQWALLKHLDTQYSD QNAFNLCFERPFVPGENRRANIQDAIWNSR KIVCLVSRHFLBGDWCLEAFYNQGRCLSDL NSALIMVVVGSLSQVQJMKHQSRGPVGKQC YLRWFBLQDDVGWHLHKLSQQLKKKEKK KDNNIPLQTVATIS 816 2166 A 6646 1 3811 RDRAGVRPAGKQHAAAAFYDVGGDRPWDS GNTQLPFRNPVKANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAGHONEFFYTAFKQ KKQQTAATGNQATPKTAPATMSTPTLLVAT AVIAYSYTNGQVVKQGKFGAAVLGHHTTR EYRLLLYISQQQPVTVARHHVIELMVPRNNY STFYDDDRQNWSIMFSESKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFOHNUGOVPDSTANKDKLLRLK LGSGRVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVICWTQATDSLLVFEVEVERRVKALA KDSGSDGHSVSSRDSAAPSPFGADNLSADPV VSPFTISFFKSGPALRTKSNSLSEQLANTSDD AVKAKLISRMAKNGGVRLLILVPDA CAVGSEGVICWTQATDSLLVFEVEVERRVKLAL KDSGSDGHSVSSRDSAAPSPFGADNLSADPV VSPFTISFFKSGPALRTKSNSLSEQLANTSDD AVKAKLISRMAKNGGVRALFLIPLDSNDSEL EDVNTLQGGGPVTFSVGSGMASFLUTBAR QHNTEIRMAVSKVADKMDHLMKVEELQKH MTSQAPQSVTGLQAPSAALMQVSLDSHSA VSGNAQSFQFYAGMQAYAYPQASAVTSQLQ PVPLYPAPLSQPPHGGSGMASFLMTBAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSMSVTURFSIMMSNIGQNER LKQELEKSNRIEGQNDKISELIERNGRYVEGS NLMMREKNNSLQTATRTNQARVLHARQEKA KVTEELAAATAQVSHJQLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLLKKTRVSTDQAAAEDIS RKSYQEELDKLRQLKKTRNSLQCRSTSLASSBORDE BERKGRRRRSTQ	1	1	ł	}	1	1	
FIPSIPDIFLSONKLYTLEKINLTANLIHLISENS LENLDLYFLLENPHOLDLANNTESCSGDO TPSENFSLEQLFLGENMLQLAWETELCWIDVE GILSHLQVLYININPLIQUENDMINSESCHOPSHITALR GLSIANSNELTV.SHNDLPANLELDISENQLL APPIDVFVSISVLDTIMERICECELSTINVL APPIDVFVSISVLDTIMERICECELSTINVL APPIDVFVSISVLDTIMERICECELSTINVL APPIDVFVSISVLDTIMERICECELSTINVL NHTINVTIAGPADIYCVYPDSLSGVSLFSLSTE GCDEEVILSKRSI.FIVCTVTLTLFLMTILLV TKRRGFCFICYKTAQRLVKNDHPGGTEPDMY KYDAYLCFSSKDTTWQWALLKHLDTQYSD QNEFNLCFERDBYVGONGLANSTVAQGRCLSDL NSALIMVVOSLSGYQLMKHOSRGFVQKQQ YLRWFEDLQDWGWFLHKLSQQLLKKEKK KDNNIPLQTVATIS 816 2166 A 6646 1 3811 RDRAGVRAGQHAAAAFDVGGDEPWDS GGARLASLFGLDQAAAGHONEFPQYTAPKOQ KKQQGTAATGNQAAMFAQAEDDTDFLSPS GGARLASLFGLDQAAAGHONEFPQYTAPKOQ KKQQGTAATGNQAAMFAQAEDDTDFLSPS GGARLASLFGLDQAAAGHONEFPQYTAPKOQ KKQQGTAATGNQAATYTAPATMSTPTILVAT AVHAYRYTNQYVKQGKFGAAVLONHTIR EYRILLYISQQPVTVARIHVNELMVRPINNY STFYDDDQRQWNSIMBESSKAAVEFNFULVAT AVHAYRYTNGQYVKQGKFGAAVLONHTIR EYRILLYISQQPYTVARIHVNELMVRPINNY STFYDDDQRQWNSIMBESSKAAVEFNFULVAT AVHAYRYTNGQYVKQGKFGAAVLONHTIR EYRILLYISQQPYTVARIHVNELMVRPINNY STFYDDDQRQWNSIMBESSKAAVEFNFULVAT AVHAYRYTNGQYVKQGKFGAAVLONHTIR EYRILLYISQQPYTVARIHVNELMVRPINNY STFYDDDQRQWNSIMBESSKAAVEFNFULVAT AVHAYRYTNGQYVKQGKFGAAVLONHTIR EYRILLYISQQPYTVARIHVNELMVRPINNY STFYDDDQRQWNSIMBESSKAAVEFNFULVARIN STFYDDQRQWNSIMBESSKAAVEFNFULVARIN STFYDDQRQWNSIMBESSKAAVEFNFULVARIN VSDRSGGGIGWTQTATDSILVFEVEYRRVLA KDSGSDGHSVSRDSAAPSPIPGADNLSADPV VSPTTSSPFKSGPBALRXSNLSEQCLANTSD AVKAKLISRNAMKMQQPMJPIPPQLDSNDSEL EDVINTLQGGGQPVYTSVSYSLQAANTSDLQ PVRPLYPAPLSQPHPQGSGDMASFLMTEAR QHNTEERMAVSIXVADKOMDHLMTKVELQKUR SAGNSMLIPSMSVTMETSMIMSNIQRIQGNRE LKQELLKSNRLBEQNDKSBLLERNQRYYEGS NAMMERRNNSLQTATRAQUERLDET QRQLTESLKSTDLLAGCLKVTSLEEGLTDL REKSTYGERTANAVERCHULGAKNOMGVSVAAAS DYSELVKIMNQVPQSLRREFELEESYNGRT LGTIMTTIMVTJCLLQLANQECKEKESSEDEE EKAEERPRRFSQQSAASSGQPQAPLARERP ESPMYPSQVELDKRULKKDNTMSTVAGERDEA KVTEELAAATQVSHLQLLKKTRVSTDQAAABOLS LQABLYCHLASAKDGLTRVLGGP TSIPPPLGGVSANDSCCESLAASPMAAKUPDN PGRKVVCWEVAPDGPLQSSTTALSTJOPE GDPLALOPESSPGEPPLGQTSTTALSTDDEE GDPLALOPESSPGEPPLGQTSTALSTDDEE GDPLALOPESSPGEPPLGQDSTALSTDDEE	1						NLQVLNLSYNLLGELYSSNFYGLPKVAYIDL
LENDLY LIVEN LOUND IN CONTROL OF THE ACT OF	İ	l					QKNHIAIIQDQTFKFLEKLQTLDLRDNALTTIH
TYSENYSLEQLFLGENMLQLAWETELCWDVY GGLSHLQVLYNINT/NISLPGYSHLTALR GLSLNSNRLTVLSHNDLPANLELDISRNQLL APPPDYVSLSVLDITHRKFICECELSTRIWL MHTNVTIAGPPADIYCVYPDSLSCVSLFSLSTT GCDEEVILKSLKFSLFIVCTVTLTLFLMTILLTY TKFRGGCFICYXTAGRUVYRDHGQOTEPDMY KYDAYLCFSSKDFTWVQNALLKHLDTQYSD QNRFNLCEERDPYGGNRFNANIQDAIWNSR KIVCLVSRHFLRDGWCLEAFSYAQGRCLSDL NSALIMVVOUSLSQVJOLMKHGSRGFVQKQQ YLRWPEDLQDVGWFLHKLSQQILKKEKKK KDNNIPLQTVATIS 816 2166 A 6646 1 3811 RDRAGVRPAGKQHAAAAFYDVGGDRPWDS GNTQLPFRNYKANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAGHONEFFOYTAFKQ KKQQGTAATGNQATPTTAPATMSTPTILVAT AVHAYXYTNGQYVGGKFGAAVLGHTTR EYRILLYISQQQPYTVARHWYELMVRPNNY STFYDDGRQNWSIMFESSKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLIRLK LGSGKVKGWEDDMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDISILVFPVEVRRVKIA KDSGSDGHSVSSRDSAAPSPFGADNISADPV VSSPTSSFFKSGPALRTKSNSLSEQLANTSPD AVKAKLISRMAKNGQMPJIPLPQLDSNDSEI EDVNTLQGGGPVYTFSVPSLQPAHPALPQ MTSQAPGPSVTGLQAFSAALMQVSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYAPLSQPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIGENER LQGELEKSNREEGOMISELERNOKRYVEGS NLMMEKRNNSLQTATENYQARVLSELQFT QMQLTESLKSTDLLRGQLTKVQAKLSELQFT QMQLTESLKSTDLLRGQLTKVQAKLSELQFT QMQLTESLKSTDLLRGQLTKVQAKLSELQFT RSSYQGELDKLRQLLKKTRVSTDQAAAGQLS PVRPLYAPLSQPHFQGSGDMASPLAMFE SCAQGSKYSEKGNRQLELKVTSLEELTDL RVEKESLEKNLSERKKSAQCRSQAEEDIDIT RSSYGELDKLRQLLKKTRVSTDQAAAGQLS DPSEKVKKIMNQVFGSLREFELEESYNGRT LGTIMNTIKMVTLQLLQQCGKERESSGEEE EKAEERRRRPSGGSASASSOQPQAPLINEERP SPSWCVKRENDQVFGSRASSESGEEE EKAEERRRRPSGGSASASSOQPQAPLINEERP SPSWCVKKIMNQVFGSLREFELEESYNGRT LGTIMNTIKMVTLQLLQNQGDEKEESSGEEE EKAEERRRRPSGGSASASSOQPQAPLINEERP SPSWCVKKIMNQVFGSLREFELEESYNGRT LGTIMNTIKMVTLQLLQNQGDEKEESSGEEE EKAEERRRRPSGGSASASSOQPQAPLINEERP SPSWCVKKIMNQVFGSLREFELEESYNGRT LGTIMNTIKMVTLQLLQNQGDEKEESSGEEE EKAEERRRRPSGGGSASASSOQPQAPLINEERP SPSWCVKKIMNQVFGSLREEFLEESYNGRT BESPECTVKKIMNQVFGSLREEFLEESYNGRT BTSPPELIGDDDDDDDGGGSTTALSTJOPEE GDPLALGPESFGEPQPPQLKKDDVTSSTOPHK ELSSTEAAGSTVAGAARTPSHSNGSRSSLSGDEE DELFKGATLKAARPRAAQPEEEDEDEDVSMKGR PPPPTLFGGDDDDDDDDDGUG	İ	1		1		1	
EGI.SHLQVI_YIMENYLINSI_PPGVPSHLTAIN GLSINSINGI_YISINDI_PANLED_IDISINQI_L APRIPDVPSLSVLDITHRKFICECELSTENUL NHTMVTLAOPPADI_YCVYDSLS_GVSLFSI_SIT GCDEEV_LISLKFSLFIVCTVTLTLFLMTILTY TKFRGFCFICYKTAQRLVFKDHPQGTEPPMY KYDAYLCFSKDFTWQNALLHRIDTQYSD QNRTNLCFERDFVPGENRPANIQDAIWNS KIVCU_SRHTLRGWCLEAFSYAQGRCLSDL NSALIMVVVGSLSQYQLMKHQSIRGFYQKQQ YLRWPEDLQDVGWFLHKLSQQILKKEKKK KDNNIPLQTVATIS BIAGVSPAGKQHAAAAYDWGGDRPWDS GRARLASLFGLDQAAAGHGNEFGYVTARKQ GKRQQTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNQQYWQGKFGAAVLGNHTIR EYRILLYISQQQVTVARHIVNFELWRPNNY STFYDDQRQNWSIMFESEKAAVENKQVCLA KCNSTSSLDAVLSQDLIVADQPAVCOSLE VAYTGWLFQNHVLGQVFDSTANKDKLLIKL LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAYGSEGVIGWTQATDSILVFEVEVRRVKLA KDSGSGGHSVSRDSAAPSPHGANLSADNISADEV VSPPTSIPFKSGEPALRTKSNSLSEQLANTSDD AVKAKLISRMAKMQOPMLPIPQLDSNDSEI EDVNTLQGGQPVVTSVQPSLQAHPALPQ MTSQAPQPSVTGLQAFSAALMQVSSLDSNSL SEQAQKSFKYRAGMAYAYPSCSANTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTERMAVSKVADKMHLMTVELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIGDRIP LQGELEKSNRIEGEDNIKISELIERNQRYVEGS NLMMEKRNNSLQTATENTQARVLHARQERET LQGULTESLKETDLLRGGLTKVQAKLSELQET SEQAQKSFKSEKONRKQLELKYTSLEGHLQYQ EVCAQRDAYQKILVGLQRESVCFALCALQ QUTALIKKELERNSKYSQAEEEIDBT RKSYGEEIDKLRGLLKKTRVSTDQAAAEQL RKSYGEEIDKLRGLLKKTRVSTDQAAAEQL QUTALTKONRGVIRKLELENNSKYGKSTAAAS DPSEKVKKIMMQVFQSLRRFELEESYNGRYCE LGGELEKSNRIEGEDNIKISELIERNQRYVEGS NLMMEKRNSLQTATENTQARVLHARQERET LQGULTESLKETDLLRGGLTKVQAKLSELQET SEQAQKSFKSEKONRKQLELKYTSLEDGHR RKSYGEEIDKLRGLLKKTRVSTDQAAAEQL QUTALTKONRGVIRKLELENNSKYGKSYAAAS DPSEKVKKIMMQVFGSLRRFELEESYNGTUL GTURNTIKMTLGLLAQUGEKSSCSEDEE EKAERPRRPGGQGAAASSGOPQAPLNEERP ESPMYSGQVVEAVEAVPLPQALTDGHRR KGDEAEALSEIKDGSLPFELSCPSYNGTUL GIDELALGFESGEPQPPQLKKDDVTSSTGPH BESELFKAGSTVAGALRSHHSDGRSSLSGOPE DELFKGATLKAALRSHARDRSSLSGOPE DELFKGATLKAALRSHARDRSSLSGOPE DELFKGATLKAALRSHARPKAQPEEBDEDEVSMKGR PPPTTLFGDDDDDDDDDUG	l		•	 			
GLSLNSRIT-IVLSRINDLPANLEILDISRNOLL APPIPDVYSLSLVDITHKRICLESTFINM. APPIPDVYSLSLVDITHKRICLESTFINM. NHTHVITIAOPPADIYCUYPDSLSGVSLFSLSTE GCDEEVLKSLKFSLFIVCTVTLTLFLIMITLY TKRGGCFICYKTAQRUYKDHPQGTEPDMY KYDAYLCFSSKDFTWVQNALLKHLDTQYSD QNRRNLCFSERDFYVQNALLKHLDTQYSD QNRRNLCFSSKDFTWVQNALLKHLDTQYSD QNRRNLCFSSKDFTWVQNALLKHLDTQYSD QNRRNLCFSSKDFTWVQNALLKHLDTQYSD QNRRNLCFSSKDFTWVQNALLKHLDTQYSD NSALIMVVYGSLSQYQLKKKEKEK KINNIPLQTVATIS 816 2166 A 6646 1 3811 RDRAGVRPAGKQHAAAAFYDVGGDRPWDS GNTQLPPRNPVKANAMFGAGDEDDTDIS-IS GNALASIFGLQAAAGHGBFFQYTARKQ KKQGTAATONQATPKTAPATMSTPTILVA AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQQPVTVARHIVNFELMVRPNIY STEYDDQRONWSIMFSSEKAAVENKQVCIA KCNSTSSLDAVLSQDLJVADGPAVEVDKLLL LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVERVKIK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVERVKIK KDSSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPPKSGEPALRITKSNSLSEQLANTSPD AVKAKLISRMKRQOPMLPIPGADNLSADPV VSPPTSIPPKSGEPALRITKSNSLSEQLANTSPD MTSGAPOPSVTGLQAPSALMGVSSLDSHSA VSGNAQSFQPYAGBAQAYAYPQASAVTSQL MTSGAPOPSVTGLQAPSALMGVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQL MTSGAPOPSVTGLQAPSALMGVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQL PVRPLYPAPLSGPPIFTQGSGDASFILMIZAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNISHLIPSMSVTMETSKIMSNIGRIGICIBER LKGELLEKSNREEDNIKISELERRQRYVEGS NLMMEKRNISLQTATENTQARVLHABGKETEL QMQLTESLKETDLLRGGLTKVQAKLSELQC PVRPLYPAPLSGPKIPGSGMASASIGNAGPHENERP SEGMGSKKSEKGNRKQLELKVTSLEEELITDL RVSKSELBKNILSERKKKSAQERSQAEEELITDL RVSKSELBKNILSERKKKSAQERSQAEEELITDL RVSKSELBKNILSERKKKSAQERSQAEEELITDL RVSKSELBKNILSERKKKSAQERSQAEEELITDL RVSKSELBKNILSERKKKSAQERSQAEEELITDL RVSKSELBKNILSERKKKSAQERSQAEEELITDL RVSKSELBKNILSERKKKSAQERSQAEELIT LGTIMNTIKMTJCLLLYCTRYSTDQAAABQL SVCAQRADYQKIQLIKKTRYSTDQAAAAQ QTTALTKQNREQHKKLELENNSQUNGSGVEAAAS DPSEKVKKIMMQVFQSLREFELESSYNGTU LGTIMNTIKMTJCLLLQCERGESSEGEE EKAEERPRRPSGGOSAASSGOPQAFLNREEP ESPMYSGQVVCBAVAPLOPQLGSSTSTGOGHE EKSELBKDGSLPPELSCPSHRVLGPP TSPPEPLLPVSMDSECEESLAASPMAAKVPDN PSGKWCVREVAPDOPQLQSSTSTSTDOFGE GDPLALGERSGEFOPPQLKKDDVTSSTGPHE SLESSFAGGSTVAGAALRSHHNSORSUSGODE DELFKGATLKALKARPKAQPEEDEDEV			f	1			
APNPDVYSLSVLDITHNKFICECELSTFINIA NHTHVTLOOPPADITVCVYPOSLSVESLSIST GCDEEVLKSLKFSLFIVCTVTLTLFLMTILTV TKFRGFCICTXTAQRLVFKDHPQGTEPDMY KYDAYLCFSKDFTWQMALLKHLDTQYSD QNRPNLCFEERDFVPQENRPANIQDAUMNSR KIVCLVSRHFLRDGWCLEAFSPANIQDAUMNSR KIVCLVSRHFLRDGWCLEAFSPANIQDAUMNSR KIVCLVSRHFLRDGWCLEAFSPANIQDAUMNSR KIVCLVSRHFLRDGWCLEAFSPANIQDAUMNSR KIVCLVSRHFLRDGWCLEAFSPANIQDAUMNSR KIVCLVSRHFLRDGWCLEAFSPANIQDAUMNSR KIVCLVSRHFLRDGWCLEAFSPANIQDAUMNSR KIVCLVSRHFLRDGWCLEAFSPANIQDAUMNSR KIVCLVSRHFLRDGWCLEAFSPANIQDAUMNSR GRALASLFGLDQAAAGHGNEFFQVTAGE GRALASLFGLDQAAAGHGNEFFQVTAGE KKQQTAATGNQATFKTAPATITLLVAT AVHAYRYTNQQYVKQGKFGAAAVLGHTTF EYRILLYISQQQVTVARHIVNFRUNYNNY STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVQBLE VAYTGWFQHVLQGVFDSTANICKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPY CAVGSEGVIGWTQATDSILVEFEVERVKVIA KDSGSDGHSVSSRDSAAPSPIPGADNISADPV VSPPTSPFKSGEPALRTSKNSLSQLANTSPD AVKAKLISRMAKMGQPMPILPPOLDSNDSEI EDVNTLQGGGQPVVTPSVQPSLQPAHFALO, MTSQAPQPSYTGLQAFSAALMQVSSLDSHSA VSGNAQSRQPYAGMQAYAYQASAVTSQLANTSPD AVKAKLISRMAKMGQPMPILPPOLDSNDSEI EDVNTLQGGGGPVVTPSVQPSLQPAHFALDA VSGNAQSRQPYAGMQAYAYQASAVTSQLANTSPD AVKAKLISRMAKMGQPMPILPPOLDSNDSEI EDVNTLQGGGPVTFSVQPSLQPAHFALDA VSGNAQSRQPYAGMQAYAYQASAVTSQLANTSPD AVKAKLISRMAKMGQPMPILPPOLDSNDSEI EDVNTLQGGGPVTFSVQPSLQPAHFALDA VSGNAQSRQPYAGMQAYAYQASAVTSQLANTSPD AVKAKLISRMAKMGQPMIPLLFPOLDSNDSEI EDVNTLQGGGPVTFSVQPSLQANTSPD AVKAKLISRMAKMGQPMIPLLFPOLDSNDSEI EDVNTLQGGGPVTFSVQPSLQANTSBHAA VSGNAQSRQPYTGQAASATHMEAR QHNTEIRMAVSKVADKMDHLMTKVELIQKH SAGNAYHAEGEAA KVTEELAAATAQVSHLQLKTKTVSTDQAAEQEA KVTEELAAATAQVSHLQLKKTRVSTDQAAEQEA KVTEELAAATAQVSHLQLKTRVSTUSTDQAAEQEA KVTEELAAATAQVSHLQLKGRCCALALQA QTTALTKQNEQHIKLEKRNSQNGGVAAAS DPSEKVKKIMQVPGGGRESSSGEGEE EKAEERPRRPSQCQSAASSGOQAPLNREEP ESPMYPSGQVVECAKCHLLASAKDEHLQQYQ FVCAQRDAYQQKLVQLOEKSTVSTDQAAAAS DPSEKVKKIMQVPGGGRESSTAAAS DPSEKVKKIMQVPGGGRESSTAASLTSDPEE GDPLALGGESFGCPPQPLKKDDVTSSTGPH VSGRVCVEVAPDQPLQESSTRLSLTSDPEE GDPLALGGESFGCPPQPQLKKDDVTSSTGPH SGSSCAGALSEIKGAALRSHHNSQRSSLSGOPE DELFRGATTKAAALRSHAPRAQFEEDEDEVSMKGR PPPTTLFGGDDDDDDDDUG					!		
NHTHVTIAOPPADIYCVYPDSLSGVSLEISGVSLEISG GOEBEVLKSLKSLFIVCTVTLILAMILTY TKPRGFCFCYKTAQRLVPKDHPQGTEPDMY KYDAYLCFSKDFTWQNALLKHLDTQYSD QNRPNLCFEERDFYPGENRPANIQDAIWNSR RIVCLVSRHFLROGWCLEAFSYAQGRCLSDL NSALIMVVVGSLSGYQLMKRIGFYOYQO YLRWPEDLQDVGWFLHKLSQQILKKEKEKK KDNNIPLQTVATIS 816 2166 A 6646 1 3811 RDRAGVRPAGKQHAAAAFYDVGGDRPWG GNTOLPPRNPVKANAMFOAGDEDDTDFLSPS GGARLASLFGLDQAAAAGHGRFFQVTARKQ KKQQGTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQQPVTVARHINNEELMVRPNNY STEYDDQRQNWSIMFBSEKAAVLGNHTTR EYRILLYISQQQPVTVARHINNEELMVRPNNY STEYDDQRQNWSIMFBSEKAAVLGNHTTR EYRILLYISQQQPVTVARHINNEELMVRPNY STEYDDQRQNWSIMFBSEKAAVLGNHTTR EYRILLYISQQQPVTVARHINNEELMVRPNY CAVGSEQVIGWTQATGGTAVACKGKGKRLLIVPPA CAVGSEQVIGWTQATGGTAVACKGKGKRLLIVPPA CAVGSEQVIGWTQATGGTAVACKGKGKRLLIVPPA CAVGSEQVIGWTQATGGTAVACKGKGKRLLIVPPA CAVGSEQVIGWTQATGGTAGATACKGKRLLIVPPA CAVGSEQVIGWTQATGGTAVACKGKGKRLLIVPPA CAVGSEQVIGWTQATGGTAVACKGKGKRLLIVPPA CAVGSEQVIGWTQATGGTAVACKGKGKRLLIVPPA CAVGSEQVIGWTQATGGTAVACKGKGKRLLIVPPA CAVGSEQVIGWTQATGGTAVACKGKGKRLLIVPPA CAVGSEQVIGWTQATGGTAVACKGKGKRLLIVPPA CAVGSEQVIGWTQATGGTAVACKGCKGKTLLIVPPA CAVGSEQVIGWTQATGGTAVACKGCKGKTLLIVPPA CAVGSEQVIGWTQATGATACKGCKGKTLLIVPPA CAVGSEQVIGWTQATGATACKGCKGKTLLIVPPA CAVGSEQVIGWTQATGATACKGCKGCKGCCCCCCCCCCCCCCCCCCCCCCCCC]			,		
GCDEEV,KSLKPSI,PIVCTVIT.IT.IHILTY TKFRGFGCYXTAGAR, LAPHOPGGTEPDMY KYDAYLCFSSKDFTWVQNALLKHLDTQYSD QNRPALCFEERDFVRGENRPANIQDAIMNSR KIVCLYSRIFLRDGWCLEAFSYAQGRCLSDL NSALIMVVVGSLSQYQLMKHQSIRGPVGCQ YLRWFEDLQDYGWFLHKLSQILKKEKEKK KDNNIPLQTVATIS 816 2166 A 6646 I 3811 RDRAGVRPAGKQHAAAAFYDVGGDPPWDS GNACLASIFGLDQAAAGHGMEFGYTARKQ KKGQTAATGNQAAAGHGMEFGYTARKQ KKGQTAATGNQAAAGHGMEFGYTARKQ KKGQTAATGNQAAAGHGMEFGYTARKQ KKGQTAATGNQAAAGHGMEFGYTARKQ KKGQTAATGNQAAAGHGMEFGYTARKQ KKGQTAATGNQAAAGHGMEFGYTARKQ KKGQTAATGNQAAAGHGMEFGYTARKQ KKGQTAATGNQAAAGHGMEFGYTARKQ KKGQTAATGNQAAAGHGMEFGYTARKQ KKGGTAATGNQATYATAATMSTETILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQQPTVARHATMSTETILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQQPTVAGHGVEGWEDW STYPUDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVAGPVEVGOSLE VAYTGWLPQNHVLGQVPDSTANKDKLLRLK LGSGKVKGWEDGMLGMKKGGKRLLIVPPA CAYGSEGVIGWTQATDSILVFEVEVKRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIFFKSGFPALRTKSNSLSEQLANTSTOP AVKAKLISRNAKMGQPMLPIPPCLISNDSEI EDVNTLQGGGQPVVTTSVQFSLQPAHFALPQ MTSQAPQPSVTGLQAFSAALMQVSSLDSNSA VSGRAQSFQPYAGMQAYATYQASAYTOSLQ PVRRLYPAHSQPHPTQGSDMASFLIMTEAR QHNTEIRMAVSKVADKMDHLMTKVELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQELEKSNRIEEQNDKISELIERNQRYVEGS NLMMEKRNISLQTATENTGANHASIQEKA VTELAAATAQVSHLQLKMTAHQKETEL QMQLTESLKETDLIRNQLTKVQAKLSSLQET SEQAQSKFRSEKQNSKQLELKVTSLEEELTDL RVEKESLEKNLSERKKSAQERSQAEEDIDEI KSSYQEELDKLRQLLKKTYSTDQAAAEQLS DYGEKVKIMNQVYGSLREFELLESTYNGRTI LGTIMNTIKMVTLQLLNQGEKEESSSEEGE EKAEERRRPSQEGSASASSGQPQAPINREP ESPMYPSGQVYGLAPQPAQLTTSODGHRR KGDSEABALSEIKDGSLPPELSCTPSHRYLGP TSIPPPLGPVQLKDDVITSODGHRR KGDSEABALSEIKDGSLPPELSCTPSHRYLGP TSIPPPLGPVQLKDDQLTTSODGHRR KGDSEABALSEIKDGSLPPELSCTPSHRYLGP TSIPPPLGPVQLKDDDLDIDIMURC ELSTEAGSTVAGAALRPSHSQRSSLSQDEE EKAEERRRPSQEGSASASSGQPQAPTREPP ESPMYPSGQVYGEAVPLPOPQLTTSODGHRR KGDSEABALSEIKDGSLPPELSCTPSHRYLGP TSIPPPLGPVAGATAGAALPDN PSGKVCVREVAPDGPLQESTRLSLTSDDFEE GDPLALGPESPGEPPPQLKKDDVTSSTGPHK ELSTEAGSTVAGAALRPSHSQRSSLSGDEE EKAEERRRPSQEGOSASASSSGQPCAATSIGPHK ELSTEAGSTVAGAALARPSHORSSLSGDEE EKAEERRRPSQEGOSASASSGQPCAPATREPP	j	1]			l
TKFRGFCFICYKTAQRLYFRDHPQGTEPDMY KYDAYLCFSERDFTWQNALLKHLDTQYSD QNRFNLCFEERDFVPGENRPANIQDAIWNSR KIVCLVSRIFTERDGWCLEAFSYAQGRCISDL NSALMYVVGSLSQYQLMKHQSIRGFVQKQC YLRWPEDLQDVGWFLHKLSQQILKKEKEKK KDNNIPLQTVATIS 816 2166 A 6646 I 3811 RDRAGYRPAGKQHAAAAFYDVGGDRPWDS GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAAGHGNEFFQYTAPKQF KKQQTTAATGNQATPKTAPATMSTFITLWAT AVHARYSTYNGQVYKQGKAAAAFVDWGGDRPWDS GGARLASLFGLDQAAAGHGNEFFQYTAPKQF KKQQTTAATGNQATPKTAPATMSTFITLWAT AVHARYSTYNGQVYKQGKAAAAFUNHTTR EYRILLYISQQQPVTVARHVYFELWWPNNY STFYDDQRQMWSIMFSSKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLPQNHVLAQVFDSTANKDKLRLK LGSGKVKKWEDDMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPANLSQDSLIVABGPAVEVSDS AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGQPVVTTSVQPSLQFAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGRAQSFGYFAGMQATAYPQASAVTSQLQ PVRRLYPAPLSQPPHFGGSGDMASFLMTEAR QHNTERMAVSKVADEMDHITK VEELQKH SAGNSMLPSMSVTMETSMIMSNIQRIQENER LKQELLEKSNRIBEQNDKISELLERNQRYVEGS NLMMEKRNNSLQTATENTGARV-HAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKSTDLLRGQLTKASLSLCBT GQAQSKFKSEKQNEKQLELKVTSLEEBLTDL RVEKESLEKNLSERKKKSQAGESQAEEEIDET RKSYQEELDKLRQLLKKTYSTOQAAAEQLS LVQAELQTQWBAKCEFILLQQYQ EVCAQRDAYQKI.VQLQEKSVCFAICLALQA QITALTKQNEQHIKLEKNKSQNGGVEAAAS DPSEKVKKIMQVFOSLREFLESSYNGGTT LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERRRPSQEQSASASSGQPQAPLTRERP ESTMYSEGVYDEAVTPLOPQLTTSQDCHRR KGDSEABALSEIKDGSLPFELSCTPSRKTLGPP TSIPPELGFVSMDSCCESSLASPMAAKUPDN PSGKVVCVREVAPPDPQLLTTSQDCHRR KGDSEABALSEIKDGSLPFELSCTPSRKTLGPP TSIPPELGFVSMDSCCESSLASPMAAKUPDN PSGKVVCVREVAPPDPQLLAKDDVTTSTGPHK ELSSTEAGSTVAGAALRESTRAKSSLSODEE EKAEERRRPSQEGSASASSGQPQAPLTTSQDCHRR KGDSEABALSEIKDGSLPFELSCTPSRKTLGPT TSIPPELGFVSMDSCCESSLASPMAAKUPDN PSGKVCVREVAPPDPQLQESSTRLSITSDPEE GDPLALGFESPPQCPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRSPHSQRSSSSGDEE EKAEERRAPRAPGQEGSASASSSGQPCAPTREPP ESTMYSEGVYGABDSCCESSLASPMAAKUPDN PSGKVCVREVAPPDPQLQESSTRLSITSDPEE GDPLALGFESPPQCPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRARFARAKSSLSGDEE DELFKGATLKALRFKAQPEEDEDEVSMKGR PPPTTIFGDDDDDDDDWLG		1	1				
KYDAYLCESSKDFTWVQNALLKHLDIQYSD QNRPALCEERPPVGCREPPVGCRAND(DAIWNSR KIVCLVSRHFLRDGWCLEAFSYAQGRCLSDL NSALIMVVVGSLSYQLMKHQSIRGFVQKQ YLRWFEDLQDVGWFLHKLSQILKKEKEKK KDNNIPLQTVATIS BIRAGVRPAGKQHAAAAFYDVGGDRPWDS GNTQLPFRNPVKANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAAGHGNEFFGYTAFKQF KKGQGTAATGNQAAYGHGMFFGYTAFKQF KKGQGTAATGNQAAYGHGMFFGYTAFKQF KKGQGTAATGNQAAYGHGMFFGYTAFKQF KKGQGTAATGNQATPKTAFTGTILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQPVTVARIHVFELMVPRNYY STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLVADGPAVEVGDSLE VAYTGWLPQNHVLGQVPTDSTANKDKLLRIK LGSGKVKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGYJGWTQATDSILVEVYRVKLA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPLIPPQLDSNDSEI EDVNTLQGGQGPVVTSVQPSLQPAHFALPQ MTSQAPQPSYTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPYAGMAYAYPQASAVTSQLQ PVRPLYPAPLSQPPYAGMAYAYPQASAVTSQLQ PVRPLYPAPLSQPPYAGMAYAYPQASAVTSQLQ PVRPLYPAPLSQPPYAGMAYAYPQASAVTSQLQ PVRPLYPAPLSQPYAGMAYAYPQASAVTSQLQ PVRPLYPAPLSQPYAGMAYAYPQASAVTSQLQ PVRLEIRMAVSKVADKMBHLMTKVEELQKH SAGNSMLPSMSVTMETSMMSNIQRIIGENER LKQELLEKSNIREGNONKISERNQRYVEGS NLMMEKRNNSLQTATENTOARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKSTDLLRGQLTKVQAKLSELQET RVSEQSKFRSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAGRSQAEEDEDEI RKSYQEELDKLRGQLKKTRVSTDQAAAGQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ FVCAQRDAYQQKLVQLQEKSVGFACLCALQA QITALTKQNEQHIKELEKNKSQMSGYEAAAS DPSEKVKKIMNQVFQSLREELESSYNGGRTU LGTIMNTIKMVTLQLLAQQCQEKEESSSEEEE EKAEERPRRPSQCQSASASSQOPQAPLNREPP ESPRVPSBQVVEEAVPLPPQLLTTSQDGHRR KGDSEAAALSEIKDGSLPPELSCHSRHYLGPP TSPPPELGFYSMGSCEESSLAASPMAAKUPDN PSGKVCKVRVAPDGPQLSKDDVTSSTGPHK ELSSTEAGSTVAGAALRSHHSQRSSLSQDEE EKAEEPRPRSQCQSASASSSQOPCAPLANREPP ESPRVPSBGVVEEAVPLPPQLATTSDGDGHRR KGDSEAALSEIKDGSLPPELSCHSRHYLGPP TSPPPLGFYSMDSECEESLAASPMAAKUPDN PSGKVCKVREVAPDGPQLRSTRJSLTSDGDEB GDPLALGPESPGPOPPQLKDDVTSSTGPHK ELSSTEAGSTVAGAATLARRFAQPEEEDEDEVSMKGR PPPTTFGDDDDDDDUDVT							
ONRFALCFEERDFYOGENRPANIQDAIWNSR KIVCLVSRIFERDGWCLEAPAGORCLSDL NSALIMVVVGSLSQYQLMKHQSIRGFVQKQQ YLRWPEDLQDVGWFHKLSQQILKKEKKK KINNIPLQTVATIS 816 2166 A 6646 1 3811 RDRAGVRPAGKQHAAAAFYDVGGDRPWDS GGARLASLFGLDQAAAGHGMEFFQYTAPKQF KKGQGTAATGNQATPKTAPATMISTILVAT AVHAYRYTNGQYVKQAKFGAAVLGNHTIR EYRLLYISQQQPVTVARIHVFELMVRPINIY STFYDDDR,GWNSIMFESSKAAVERNKQVCLA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTOWLPOHIVLGQVPSTANKDKLLRK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVVEXRVIA KDSGSDGHSVSSRDSAASPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLANTSED AVKAKLISRMAKMGQPALPIPPQLDSINDSEL EDVNTLQGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAMTSVPQASAVTSQLQ PVRRLYPAPLSQPPHPGGSGDMASFLMTEAR QHNTEIRMAVSVADKMDHTKVELQKH SAGNSMLPSMSVTNETSMIMSNIGRIGENER LKQELLEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRINSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKSTDLLRGGLTKAALSSQLEELTDL RVEKESLEKINLSERKKKSAGNSQLEELDEIT RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKI.VQLQEELESYYNGRTI LGTIMMTIKMVTLQLLNQGQEKEESSSEEEE EKAEERPRRPSQGSASASSGQPQAPLINERP ESPMYPSGVVEEAVPLPQALTTSQDCHRR KGDSEAAALSEKOSLSPELSCTYNGRTI LGTIMMTIKMVTLQLLNQGQEKEESSSEEEE EKAEERPRRPSQGSASASSGQPQAPLNTERP ESPMYPSGVVEEAVPLPQALTTSQDCHRR KGDSEAAALSEKKOSLSPELSCTYNGRTI LGTIMMTIKMVTLQLLNQGQEKEESSSEEEE EKAEERPRRPSQGSASASSGQPQAPLNTERP ESPMYPSEGVVEEAVPLPQALTTSQDCHRR KGDSEAAALSEKKOSLSPELSCTYNGRTI LGTIMMTIKMVTLQLLNQGQEKEESSSEEEE EKAEERPRRPSQGSASASSGQPAPLNTERP ESPMYPSEGVVEEAVPLPQALTTSQDCHRR KGDSEAAALSEKKOSLSPELSCTYNGRTI LGTIMMTIKMVTLQLLNQGQEKEESSSEEEE EKAEERPRRPSQGSASASSGQPAPLTREAP ESPMYPSEGVVEEAVPLPQALTTSQDCHRR KGDSEAAALSEKKOSLSPELSCTYNGRTI LGTIMMTIKMVTLQLLNQGQEKEESSSEEEE EKAEERPRPSPGGSASASSGQPAPLTREAP ESPMYPSEGVVEEAVPLPQALTTSQDCHRR KGDSEAAALSEKKOMSUPEALTSGDCHRR KGDSEAAALSEKKOMSUPEALTSGDCHRR ESSTEAGSTVAGAALRSSTRAKKSCRSSSGDE EKAEERPRSPGGGSASASSGQPAPLTSGDCHRR ESSTEAGSTVAGAALAKRAKAQPEEDEDEVSMKGR PPPTFIFGDDDDDDDUMLCA							
KIVCLVSRHFILROWCLEARSYAGGRCLSDL NSALIMVVGSLSQVJLKKGSRRGVGKOC YLRWPEDLQDVGWFLHKLSQQLKKEKKK KDNNIPLQTVATIS RDRAGVRPAGKQHAAAAFYDVGGDRPWDS GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAAGNEDFPQYTAPKQF KKQQTAATGNQATPKTAPATMSTFTILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTIR EYRILL YISQQPVTVARUOFGAVLGNHTTIR EYRILL YISQQPVTVARUOFGAVLGNHTTIR EYRILL YISQQPVTVARUOFGAVLGNHTTIR EYRILL YISQQPVTVARUOFGAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATTDSLVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPFTSIPFESGGPALRINSLSEQLANTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSE EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ MTSQAPGSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASATSQLQ PVRPLYPAPLSOPPHFEGGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIGRIQENER LKQEILEKSKRIEEQNDKISELERNORYVEQS NLMMEKRNISLQTATENTQARVLAEQEKA KVTELLAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRQQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEELTDLL RKSYQEELDKLRQLIKKTRVSTDQAAAEQLS LVQAELQTOWEAKCEHLAAAAQSHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFAYLAALQA QITALTKQNEQHIKLLEKNKSQMSVGWAAAS DPSKEVKKIKMINQVFQSLKSMSVEWAAAS DPSKEVKKIKMINQVFQSLKRSFEELEESYNGRT LGTIMNTIKMYTLQLLNQGEQEKEESSSEIEGE EKAEEPRREPSGEVSAASSGOPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHR KGDSEAEALISEIKDGSLPPELSCIPSIRVLGPP TSIPPEPLGPYSMDSECESAASSGOPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHR KGDSEAEALISEIKDGSLPPELSCIPSIRVLGPP TSIPPEPLGPYSMDSECESAASSMOPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHR KGDSEAEALISEIKDGSLPPELSCIPSIRVLGPP TSIPPEPLGPYSMDSSCESAASSGOPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHR KGDSEAEALISEIKDGSLPPELSCIPSIRVLGPP TSIPPEPLGPYSMDSSCESAASSGOPQAPLNRERP ESSTEAGSTVAGAALRPSHHSQRSSSSGOEDE EKAEEPREPSGEPOSAASSGOPQAPLNRERP ESSTEAGSTVAGAALRPSHHSQRSSSSGOEDE EKAEEPREPSGEPOSAASSGOPQAPLNERP ESSTEAGSTVAGAALRPSHHSQRSSSSGOEDE EKAEEPREPSGEPOSAASSGOPQAPLNERP ELSSTEAGSTVAGAALRPSHHSQRSSSSGOEDE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPTTPLFGDDDDDDDD	ŀ	•					
NSALIMVVVGSLSQVQLMKHQSIRGFVQKQQ YLRWPEDLQDVGWFLHKLSQQILKKKEKKK KDNNIPLQTVATIS RDRAGVRPAGKQHAAAFYDVGGDRPWDS GGARLASLFGLDQAAAGHGNEFFQYTAPKQF KKGQGTAATGNQATKTAPATNIFTLLVAT AVHAYRYTNGQYVKQGFGAAVLGNHTTR EVRILLVISQQQPVTVARIHVNFELMVRPNNY STFYDDQRQNWSIMFBEAKAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKLA KDSGSDGINSVSRDSARAVEFNKQDLINTSPD AVKAKLISRMAKMGQPM.PILPPQLDSINDSEL EDVNTLQGGGVPVTPSVPSLQPAHFALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAAVTSQLQ PVRLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIGENER LKQELEKSNRIEEQNDKISELIERNQRYVEQS NLMMEERNINSLGTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAALSELGET RKSYQEELDKLRQLLKKTRYSTDQAAAEQLS LVQABLQTQWEAKCEHLLASAKDEHLQQYQ FVCAQRDAYQQKI.VQLQEKSVCFAACLALQA QITALTKQNEQHIKELEKNISQMSGVEAAAS DPSKKVKKIMNQVPGSLREFFELESYMGRTI LGTIMNTIKMVTLQLLNQQEQEKESSEEEEE EKAEERRRERPSGGSASASSGOPQAPLMRERP ESPMYPSEQVVEEAVPLPQALTTSQDGHRR KGDBEABALSEKIGGSLPPELSCIPIRLVLGPP FSGMVCVREVAPDGPLQESSTRLSTSDPEE GDPLALGPESPGEPQPPPQLKDDVTSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DGLFKGATLKALRKARPSGRSSLSGOEE DGLFKGATLKALRKARPSGRSSLSGOEE DGLFKGATLKALRKARPSGRSSLSGOEE DGLFKGATLKALRKARPSGRSSLSGOEE DGLFKGATLKALRKARPSGRSSLSGOEE DGLFKGATLKALRKARPSGRSSLSGOEE DGLFKKGATLKALRKARPSGRSSLSGOEE DGLFKKGATLKALRKARPSGRSSLSGOEE DGLFKGATLKALRKARPSGRSSLSGOEE DGLFKKGATLKALRKARPSGRSSLSGOEE DGLFKKGATLKALRKARPSGRSSLSGOEE DGLFKKGATLKALRKGRSSLSGOEE DGLFKKGATLKALRKARPSGRSSLSGOEE DGLFKKGATLKALRKGRSSLSGOEE DGLFKKGATLKALRKGRSSLSGOEE DGLFKKGATLKALRKGRSSLSGOEE DGLFKKGATLKALRKGRSSLSGOEE DGLFKKGATLKALRKGRSSLSGOEE DGLFKKGATLKALRKARPGREEEDEDEVSMKGR PPFTPLFGDDDDDDDDURG				1			
VILRWPEDLQDVGWPLHKLSQQILKKEKEKK KUNNIPLQTVATIS							
KDNNIPLQTVATIS	l .					,	
GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS GGARLASLFGLDQAAAGHGNFFQYTAPKQF KKGQGTAATGNQATPKTAPATMSTPTLWAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQQPVTVARHIVNFELMVRPNNY STFYDDDQRONWSIMFSEKAAVEFRKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LOSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKLIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGGPMLPIPQLDSNDSEI EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHGOSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIQENER LKQEILEKSNRIEEQNKISELIERNGRYVEQS NLMMEKRNNSLQTATETOARVLHAEQEKA KVTEELAAATAQVSHLQLKKVTAHQKKETEL QMQLTESLKSTDLERONKISELIERNGRYVEGS NLMMEKRNSLQTATETOARVLHAEQEKA KVTELAAATAQVSHLQLKKVTAHQKKETEL QMQLTESLKSTDLERONKISELIERNGRYVEGS SEQAQSKFKSEKQNRKQLELKVTSLEEELIDDL RVEKESLEKNLERGKLKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELIDDL RVEKESLEKNLERGLKKVQAKLSELQET GEQAGNAYQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ FVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKLLKKNSQMSGVEAAAS DPSEKVKKIMNQVFQSLREEFLEESTNIGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSEEEE EKAEERPRPSQEQSAASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAALSEIKOSISPPELSCIPSTRYLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKPDN PSGKWCVREVAPDGPIQESSTIRSLT.SUPFEE GGPLALGFESGGEPPPLLKCPDTTSTGPHK ELSSTEAGSTVAGAALRPSHISQRSSLSGDEE DBLFKGATLKALRPFKAQPEEEDEDEVSMKGR PPPPTLEGDDDDDDDDUGG							
GGARLASLFGLDQAAAGHGNEFFQYTAPKOP KKGQTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQYVKQGKGAAVLGNHTITE EYRILLYISQQQPVTVARIHVNFELMVRPNNY STFYDDQRQNWSHFSSKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRVKIA KDSGSDGHSVSSRDSAAPSPIPGADDILSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQFSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQYAGMQAYAYQASAVTSQLQ PVRRLYPAPLSQPPHPQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIGNER LQEBLEKSNRIEGROKISELIERNQRYVEGS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLAGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAGERSQAEEIDET RKSYQEELDKLIRQGLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAGERSQAEEIDET RKSYQEELDKLIRQGLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL QMQLTESLKETDLKHLLEKNKSQMSGVEAAAS DFSEKVKIMNOVPGSLREFFELEESYNGRTI LGTIMTIKMVTLQLLKQEGEKEESSEEEE EKAEERPRRPSQEQSASASSGGPQAPLNEERP ESSPMYPSGQVEEAVPLPPQALTTSQDGHRR KGDSEABALSEIKEDGSLPPELSCTPSIRVTLGPP ESSPMYPSGQVEEAVPLPPQALTTSQDGHRR KGDSEABALSEIKEDGSLPPELSCTPSIRVTLGPP TSIPPPLGPVSMDSECEESLAASPMAAKPDN PSGKWCVREVAPDGPLQESSTRUSLTSDDFEE GGPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRRKAQPEEDEDEVSMKGR	816	2166	A	6646	1	3811	
KKGQGTAATGNQATPKTAPATMSTPTILVAT AVHAYRYTNGQYVKQGKFGAAVLGNHTTR EYRILLYISQQPVVTVARIHVNFELMVRPNNY STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFONHVLGQVFDSTANKDKLIRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGGEGVIGWTQATDSILVFEVEVRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPFTSIPFKSGEPALRTKSNSLSQLAINTSFD AVKAKLISRMAKMGQPMJPLIPPQLDSNDSEI EDVNTLQGGQPVVTPSVQPSLQFAHPALPQ MTSQAPQFSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPVAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHMTKVEELQKH SAGNSMLPBMSVTMETSMIMSNIGRIJGENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NILMMEKRNSLQTATENTQAVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKEFELL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSEKSLEKVSAGERSQAEEGIDET RKSYQEELDKLRQLLKKTSTDQAAAAEQLS LVQAELQTQWEAKCEHILASAKDEHLQQYQ EVCAQRDAYQQKI VQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNVQVGSLRREFELEESYNGRTI LGTIMMTIKMVTLQLLNQCQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEGVVEEAVPPQALTTSQDGHRR KGDSEABALSEIKDGSLPPELSCIPSTRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKPDN PSGKWCVREVAPDGPLQESSTRLSLTSDDFE GDPLALGPESPEGPPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DDLFKKATKARTKARTARCHSPEEDEEDEVSMKGR							GNTQLPPRNPVKANAMFGAGDEDDTDFLSPS
AVHAYRYTINGYYKQGKFGAAVLGNIHTTR EYRILYISQQQPYTVARIHVNFELMVRPNNY STFYDDQRQNWSIMFESEKAAVEFKKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVEVEVEVRRVKIA KDSGSGGHSVSSRDSAAPSPIPGADNLSADPV VSPGSGGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPLIPPQLDSNDSEI EDVNTLQGGGGPVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQYAGMQAYAPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEDONDKISELIERNQRYVEQS NILMMEKRNSLQTATENTQAVVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKONRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDBI RKSYQEELDKLRGLLKKTRVSTDQAAAEQLS LVQAELQTOWEAKCHILLASAKDEHLQQYQ EVCAQRDAYQQKLVQLOEKSVCPACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVPQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEGKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESSMVPSEGVVEEAVPLPPQALTTSQDGHRR KGDSEAALSEIKGGSLPPELSCIPSIRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKPDN PSGKNCVREVAPDGPLQESTRUSLTSDDFE GGPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR							GGARLASLFGLDQAAAGHGNEFFQYTAPKQP
EYRILLYISQQPVTVARIHVNFELMVRPNNY STFYDDQRQNWSIMFESEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQFMLPILPPQLDSNDSEI EDVNTLQGGQPVVTPSVPSLQPALPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSNSVTMETSMMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEDS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKSAQERSQAEEEIDEI RKSYQEELDKLRGLLKSKRUELLAQVG EVCAQRDAYQQKLVQLQEKSVCFAKCLALQA QITALTKQNEQHIKELERNKSOMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGOPQAPLNRERP ESPMVPSEVVEAVPLQLLKGESSCEEEL EKAEERPRRPSQEQSASASSGOPQAPLNRERP ESPMVPSELSKINSLSTLSTSDPEE GDPLALGPESSGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRSHHSORSSLSGDEE DELFKGATLKALRPKAQPEEDEDEVSMKGR							
STFYDDQRQNWSIMFBSEKAAVEFNKQVCIA KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQhIVLGQVFDSTANKDKLIRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRIKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGQVVTPSVVPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMMSNIQRIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKSAQERSQAEEIDEI' RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSVBAAAS DPSEKVKKIMOYFQSLREFFELEESYNGRTI LGTIMNTIKMVTLQLLNQEQEKEESSSEEEE EKAEERRRSPGGCSAASASGQPQAPLNRERP ESPMYPSEQVVEQAVPDGPLQESTRLSLTSUPFEE GDPLALGPESPGEPQPPQLKKDDVTSSTOPHK ELSSTEAGSTVAGAALRSHSQRSSLSGDEE DDLALGPESPGEPQPPQLKKDDVTSSTOPHK ELSSTEAGSTVAGAALRSHSQRSSLSGDEE					· .		
KCNSTSSLDAVLSQDLIVADGPAVEVGDSLE VAYTGWLFQNHVLGQVFDSTANKKLGKKLLKLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKLA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGPALRIKSNSLSEQLANTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGQPVVTPSVPSLQPALPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVELQKH SAGNSMLPSNYTMETSMIMSNIQRIIGENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMERRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKSAQERSQAEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSOMGSVEAAAS DPSEKVKKIMNQVFQSLREFFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAALNERPP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKIPDN PSGKWCVREVAPDOPLQESSTRJSLTSWDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTOPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEEDEVSMKGR PPPTPLFGDDDDDDDDWLG]						
VAYTGWLFONHVLGQVFDSTANKDKLLRLK LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVREYEVKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSDFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHPGGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSKLIPSMSVTMETSMIMSNIQRIIQENER EKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRINSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETIDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKI, VOLQEKSVCFAICLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVYEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECESLAASPMAAKIPDN PSGKVCVREVAPDGPLQESSTRLSLTSIDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHSQRSSLSGDEE DDELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFEGDDDDDDDDDWLG							
LGSGKVIKGWEDGMLGMKKGGKRLLIVPPA CAVGSEGVIGWTQATDSILVFEVEVRRVKIA KDSGSDGHSVSSRDSAAPSPIPGADNI SADPV VSPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFYGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTELLAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNKQLELKVTSLEEELTDLL RVEKESLEKNLSERKKKSAQERSQAEEEIDET RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKI.VQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEPPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPELGPVSMDSECESLAASPMAAKPDDN PSGKWCVREVAPDGPLQESSTRLSLTSDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG					•		
CAYGSEGVIGWTOATDSILVFEVEVRRYKIA KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVTFSVQPSLQPAHFALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTTAPKKETEL QMQLTESLKETDLLRQQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RSKYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFAICLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKIPDN PSGKVVCVREVAPDGPLQESSTRLSLTSUPPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DDELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDUWLG	·			i			
KDSGSDGHSVSSRDSAAPSPIPGADNLSADPV VSPPTSIFFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGPVVTPSVQPSLQPAFFALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQSAVTSQLQ PVPRLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NILMMEKRINSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQI.QEKSVCFANCLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVCKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQGEQEKEESSSEEE EKAEERPRRPSQEQSASSGOPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAFALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKPDN PSGKVCVREVAPDOPLQESSTRLSLTSUPFEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DDLFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPTLFGDDDDDDDDDWLG							
VSPPTSIPFKSGEPALRTKSNSLSEQLAINTSPD AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKI.VQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEGEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKLGGSLPPELSCIPSRIRVLGPP TSIPPELGPVSMDSECEESLASPMAAKYPDN PSGKIVCVREVAPDGPLQESSTRLSLTSUPPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRRSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR	}						
AVKAKLISRMAKMGQPMLPILPPQLDSNDSEI EDVNTLQGGGQPVVTPSVQPSLQPAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQELLEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRYSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA/CLALQA QITALTKQNEQHIKFLEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLREFFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSEEEE EKAEERPRRPSQEQSASASGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIGRVCVREVAPDGPLQESSTRLSLTSUPPE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRFSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR							
EDVNTLQGGGQPVVTPSVQPSLQFAHPALPQ MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLQQQQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESFMYPSQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSPPPEPLGFVSMDSECEESLAASPMAAKYPDN PSGK\VCVREVAPDGPLQESSTRLSLTSUPPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEDEDEVSMKGR							
MTSQAPQPSVTGLQAPSAALMQVSSLDSHSA VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNISLQTATENTQARVLHAEQEKA KVTEELAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDLL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKI.VQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKPDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHASQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG						İ	
VSGNAQSFQPYAGMQAYAYPQASAVTSQLQ PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKI.VQLQEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKYPDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDWLG							MTSOAPOPSVTGLOAPSAALMOVSSLDSHSA
PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NILMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDWLG							VSGNAQSFQPYAGMQAYAYPQASAVTSOLO
QHNTEIRMAVSKVADKMDHLMTKVEELQKH SAGNSMLIPSMSVTMETSMIMSNIQRIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKLVQLQEKSVCFACLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEEVSMKGR PPPTPLFGDDDDDDDDWLG							PVRPLYPAPLSQPPHFQGSGDMASFLMTEAR
SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER LKQEILEKSNRIEEQNDKISELIERNQRYVEQS NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKI.VQI.QEKSVCFAICLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGOPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDIDWLG					i	ĺ	QHNTEIRMAVSKVADKMDHLMTKVEELQKH
NLMMEKRNNSLQTATENTQARVLHAEQEKA KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKI.VQI.QEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG							SAGNSMLIPSMSVTMETSMIMSNIQRIIQENER
KVTEELAAATAQVSHLQLKMTAHQKKETEL QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHILLASAKDEHLQQYQ EVCAQRDAYQQKLVQI.QEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMYPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDWLG					ļ		
QMQLTESLKETDLLRGQLTKVQAKLSELQET SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKI.VQI.QEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKYPDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDDWLG							
SEQAQSKFKSEKQNRKQLELKVTSLEEELTDL RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHILASAKDEHLQQYQ EVCAQRDAYQQKLVQI.QEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\CVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHE ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDIDWLG	[l		İ	- 1	į	KVILELAAATAQVSHLQLKMTAHQKKETEL
RVEKESLEKNLSERKKKSAQERSQAEEEIDEI RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKI.VQI.QEKSVCFAVCLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKYPDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDUWLG					ı		
RKSYQEELDKLRQLLKKTRVSTDQAAAEQLS LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKI.VQI.QEKSVCFAICLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKIPDN PSGKIVCVREVAPDGPLQESSTRLSLTSIDPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHPSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDWLG		1			<u>.</u> .		DVEVEC EVALUED VERSON EBOOMERS
LVQAELQTQWEAKCEHLLASAKDEHLQQYQ EVCAQRDAYQQKI.VQI.QEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDIDWLG		1		1		į	
EVCAQRDAYQQKI.VQI.QEKSVCFA\CLALQA QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESSPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDIDWLG		1			, 1		
QITALTKQNEQHIKELEKNKSQMSGVEAAAS DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDIWLG		ļ			-		EACAORDAAOURI AUT UERRACEUTÓATOT
DPSEKVKKIMNQVFQSLRREFELEESYNGRTI LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDM PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG					ł		
LGTIMNTIKMVTLQLLNQQEQEKEESSSEEEE EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDIDWLG		1			ļ		
EKAEERPRRPSQEQSASASSGQPQAPLNRERP ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAKYPDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDIDWLG			i		ł		
ESPMVPSEQVVEEAVPLPPQALTTSQDGHRR KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHE ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDIDWLG		l			1		EKAEERPRRPSOEOSASASSGOPOAPI NDEDD
KGDSEAEALSEIKDGSLPPELSCIPSHRVLGPP TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEEVSMKGR PPPTPLFGDDDDDDDWLG			l		1		
TSIPPEPLGPVSMDSECEESLAASPMAAK\PDN PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDIDWLG		j	i	[į		KGDSEAEALSEIKDGSLPPELSCIPSHRVI.GPP
PSGK\VCVREVAPDGPLQESSTRLSLTS\DPEE GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG		,		[}		
GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG		ļ			ľ		
ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG		ł			- [GDPLALGPESPGEPQPPQLKKDDVTSSTGPHK
DELFKGATLKALRPKAQPEEEDEDEVSMKGR PPPTPLFGDDDDDDDDDWLG		į			İ		ELSSTEAGSTVAGAALRPSHHSQRSSLSGDEE
		ļ			1	1	DELFKGATLKALRPKAQPEEEDEDEVSMKGR
81/ 1216/ A 16649 163 11073 FEDEREDNICEDIDOVE/HETPALICEDIA/OTEC							
TINGSSDINGSTRY I DISTANÇOF VMGLEG	817	2167	<u>A</u>]	6649	63	1073	FFRSSSDNGSPIRQYE/HSTPAHQGPVMGLEG

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion KS/ARNSQLRIVLVGKTGAGKSATGNSILGRK VFHSGTAAKSITKKCEKRSSSWKETELVVVD TPGIFDTEVPNAETSKEIRCILLTSPGPHALLL VVPLGRYTEEEHKATEKILKMFGERARSFMIL IFTRKDDLGDTNLHDYLREAPEDIQDLMDIFG DRYCALNNKATGAEQEAQRAQILGIJQRVV RENKEGCYTNRMYQRAEEEIQKQTQAMQEL HRVELEREKARIREEYEEKIRKLEDKVEQEKR KKQMEKKLAEQEAHYAVRQQRARTEVESKD GILELIMTALQIASFILLRLFAED
818	2168	A	6660	357	1890	APSGSWTRVVLTLDPCSLRSRSPRSLLDPGMP GISARGLSHEGRKQLAVNLTRVLALYRSILDA YIIEFFYTDNLWDTLPCSWQEALDGLKPPQLA TMLLGMPGEGEVVRYRSVWPLTLLALKSTA CALAFTRMPGFQTPSEFLENPSQSSRLTAPFR KHVRPKKQHEIRRLGELVKKLSDFT/GLHPGC RRGLRPG/HLSRFMALGLGLMVKSIEGDQRL VERAQRLDQELLQALEKEEKRNPQVVQTSPR HSPHHVVRWVDPTALCELLLPLENPCQGRA RLLLTGLHACG\DLSVALLRHFSCCPEVVALA SVGCCYMKLSDPGGYPLSQWVAGLPGYELP YRLREGACHALEEYAERLQKAGPGLRTHCY RAALETVIRRARPELRRPGVQGIPRVHELKIEE YVQRGLQRVGLDPQLPLNLAALQAHLAQEN RVVAFFSLALLLAPLVETLILLDRLLYLQEQA LSP\GFHAELLPIFSPELSPRNLVLVATKMPLG QALSV\LETEDS
819	2169	A	6661	65-	2686	SGSCHCLAEASMGPWGWKLRWTVALLLA AAGTAVGDRCERNEFQCQDGKCISYKWVCD GSAECQDGSDESQETCLSVTCKSGDFSCGGR VNRCIPQFWRCDGQVDCDNGSDEQGCPPKTC SQDEFRCHDGKCISRGFVCDSDRDCLDGSDE ASCPVLTCGPASFQCNSSTCIPQLWACDNDPD CEDGSDEWPQRCRGLYVFQGDSSPCSAFEFH CLSGECIHSSWRCDGGPDCKDKSDEENCAVA TCRPDEFQCSDGNCIHGSRQCDREYDCKDMS DEVGCVNVTLCEGPNKFKCHSGECITLDKVC NMARDCRDWSDEPIKECGTNECLDNNGGCS HVCNDLKIGYECLCPDGFQLVAQRRCEDIDE CQDPDTCSQLCVNLEGGYKCQCEGFQLDPH TKACKAVGSIAYLFFTNRHEVRKMTLDRSEY TSLIPNLRNVVALDTEVASNRIYWSDLSQRMI CSTQLDRAHGVSSYDTVISRDIQAPDGLAVD WIHSNIYWTDSVLGTVSVADTKGVKRKTLFR ENGSKPRAIVVDPVHGFMYWTDWGTPAKIK KGGLNGVDIYSLVTENIQWPNGITLDLLSGRL YWVDSKLHSISSIDVNGGNRKTILEDEKRLAH PFSLAVFEDKVFWTDIINEAIFSANRLTGSDV NLLAENLLSPEDMVLFHNLTQPRGVNWCERT TLSNGGCQYLCLPAPQINPHSPKFTCACPDGM LLARDMRSCLTEG\EAAVATQETSTVRLKVS STAVRTQHTTTRPVPDTSRLPGATPGLTTVEI VTMSHQALGDVAGGRGNEKKPSSVRALSVL PIVLLVFLCLGVFLLWKNWRLKNINSINFDNP VYQKTTEDEVHICHNQDGYSYPSRQMVSLED DVA
820	2170	A	6666	17	4146	ERGISSQIKGMKSGSGGGSPTSLWGLLFLSAA LSLWPTSGEICGPGIDIRNDYQQLKRLENCTVI EGYLHILLISKAEDYRSYRFFKLTVITEYLLLF RVAGLESLGDLFPNLTVIRGWKLFYNYALVIF

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			7.4	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	1	l	}	peptide	scquence	/=possible nucleotide deletion, \=possible
		ł		sequence		nucleotide insertion
	 	 		Sequence	 	
	1		i l			EMTNLKDIGLYNLRNITRG\AIRIEKNADLCYL
ļ		Į.			J	STVDWSLILDAVSNNYIVGNKPPKECGDLCP
1		İ	1			GTMEEKPMCEKTTINNEYNYRCWTTNRCQK
İ	I					MCPSTCGKRACTENNECCHPECLGSCSAPDN
1]	ļ	1 .		}	DTACVACRHYYYAGVCVPACPPNTYRFEGW
						RCVDRDFCANILSAESSDSEGFVIHDGECMQE
	1				}	CPSGFIRNGSQSMYCIPCEGPCPKVCEEEKKT
	1	1	! !		1	KTIDSVTSAQMLQGCTIFKGNLLINIRRGNNIA
	1		i			SELENFMGLIEVVTGYVKIRHSHALVSLSFLK
i						NLRLILGEEQLEGNYSFYVLDNQNLQQLWD
ł	1	ł			ļ	WDHRNLTIKAGKMYFAFNPKLCVSEIYRMEE
}	j	l				VTGTKGRQSKGDINTRNNGERASCESDVLHF
		l	i 1			TSTTTSKNRIIITWHRYRPPDYRDLISFTVYYK
}		ŀ			}	EAPFKNVTEYDGQDACGSNSWNMVDVDLPP
1	1	[NKDVEPGILLHGLKPWTQYAVYVKAVTLTM
						VENDHIRGAKSEILYIRTNASVPSIPLDVLSAS
1		1				NSSSQLIVKWNPPSLPNGNLSYYIVRWQRQP
						QDGYLYRHNYCSKDKIPIRKYADGTIDIEEVT
						ENPKTEVCGGEKGPCCACPKTEAEKQAEKEE
						AEYRKVFENFLHNSIFVPRPERKRRDVMQVA
]						NTTMSSRSRNTTAADTYNITDPEELETEYPFF
}	1 1		Į į			ESRVDNKERTVISNLRPFTLYRIDIHSCNHEAE
]]					KLGCSASNFVFARTMPAEGADDIPGPVTWEP
]			•		RPENSIFLKWPEPENPNGLILMYEIKYGSQVE
]					DQRECVSRQEYRKYGGAKLNRLNPGNYTARI
						QATSLSGNGSWTDPVFFYVQAKRYENFIHLII
	[[ALPVAVLLIVGGLVIMLYVFHRKRNNSRLGN
			İ			GVLYASVNPEYFSAADVYVPDEWEVAREKIT
j]]		ļ	·]		MSRELGQGSFGMVYEGVAKGVVKDEPETRV
			ļ			AIKTVNEAASMRERIEFLNEASVMKEFNCHH
						VVRLLGVVSQGQPTLVIMELMTRGDLKSYLR
1			J			SLRPEMENNPVLAPPSLSKMIQMAGEIADGM
						AYLNANKFVHRDLAARNCMVAEDFTVKIGD
1	ĺ	- 1		1		FGMTRDIYETDYYRKGGKGLLPVRWMSPESL
						KDGVFTTYSDVWSFGVVLWEIATLAEQPYQ
1	1	- 1				GLSNEQVLRFV\MEGGLLDKPDNCPDMLFEL
			1		l	MRMCWQYNPKMRPSFLEIISSIKEEMEPGFRE
						VSFYYSEENKLPEPEELDLEPENMESVPLDPS
į į		i				ASSSSLPLPDRHSGHKAENGPGPGVLVLRASF
						DERQPYAHMNGGRKNERALPLPQSSTC
821	2171	A	6691	106	825	GRVLFRGCGVGHKGQVLMGTFILAQDWLSE
		ļ	1		İ	SNHVFCVSSMLRLQKRLASSVLRCGKKKVW
		- 1	- 1	[[LDPNETNEIANANSRQQIRKLIKDGLIIRKPVT
						VHSRARCRKNTLARRKGRHMGIGKRKGTAN -
		J	ŀ	1	J	ARMPEKVTWMRRMRILRRLLRRYRES/KRYR
		-		İ]	ESKKIDRHMYHSLYLKVKGNVFKNKRILMEH
	}	1	}	j	į	IHKLKADKARKKLLADQAEARRSKTKEARK
1		l	.]	1	ĺ	RREERLQAKKEEIIKTLSKEEETKK
822	2172	A	6715	772	21	DFRPGLLLPRKKKMFGFHKPKMYRSIEGC\CI
		- 1	ļ	-		SGAKSSSS\RFTDSKRYEK\DFO\SCFGLHETR\
1 1		- 1	1	J	J	SGDI/CNA/CVLL/LKRWKKLPAGSKK/NWNH
			1	ļ	l	
		- 1	- 1	i		VVDARAGPS\LKTTLKPKKVKTL\SGNRIK\ST
			1	į	l	QISKLQKEFKR\HNSDAHSTTS\SASP\AQSPLF
				ľ	l	TVNQFRWTGSDTGVGFPGSNRNHPVFSFLDL\
[[İ	- 1		ł	1	TYWKRQKICCGNYKGRFGEVLIDTHLFKPCC
922	2172	,	6727		1062	SNKKA\AAEKPEEQGPEPLPISTQEWVTEVFM
823	2173	Ą	6727	3	4063	PYLATLQLDSSLLIPPKYQTPPAAAQGQATPG
	ľ	{	- 1	ľ	ļ	NAGPLAPNGSAAPPAGSAFNPTSNSSSTNPAA
	-	i	1	ŧ		SSSASGSSVPPVSSSASAPGISQISTTSSSGFSGS
L		l				VGGQNPSTGGISADRTQGNIGCGGDTDPGQS
						

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
				sequence		SSQPSQDGQESNVPSVGSLADPDYLNTPQMN TPVTLNSAAPASNSGAGVLPSPATPRFSVPTP RTPRTPRTPRGGGTASGQGSVKYDSTDQGSP ASTPSTTRPLNSVEPATMQPIPEAHSLYVTLIL SDSVMNIFKDRNFDSCCICACNMNIKADVG LYIPDSSNEDQYRCTCGFSAIMNKLGYNSGL FLEDELDIFGKNSDIGQAAERRLMMCQSTFL PQVEGTKKPQEPPISLLLLQNQHTQPFASLN FLDYISSNNRQTLPCVSWSYDRVQADNNDY WTECFNALEQGRQYVDNPTGGKVDEALVRS ATVHSWPHSNVLDISMLSSQDVVRMLLSLQP FLQDAIQKKRTGRTWENIQHVQGPLTWQQFH KMAGRGTYGSEESPEPLPIPTLLVGYDKDFLT ISPFSLPFWERLLLDPYGGHRDVAYIVVCPEN EALLEGAKTFFRDLSAVYEMCRLGQHKPICK VLRDGIMRVGKTVAQKLTDELVSEWFNQPW SGEENDNHSRLKLYAQVCRHHLAPYLATLQL DSSLLIPPKYQTPPAAAQGQATPGNAGPLAPN GSAAPPAGSAFNPTSNSSSTNPAASSSASGSSV PPVSSSASAPGISQISTTSSSGFSGSVGGQNPST GGISADRTQGNIGCGGDTDPGQSSSQPSQDG QESVTERERIGIPTEPDSADSHAHPPAVVIYM VDPFTYAAEEDSTSGNFWLLSLMRCYTEMLD NLPEHMRNSFILQIVPCQYMLQTMKDEQVFY IQYLKSMAFSVYCQCRRPLPTQIHIKSLTGFGP AASIEMTLKNPERPSPIQLYSPPFILAPIKDKQT ELGETFGEASQKYNVLFVGYCLSHDQRWLL ASCTDLHGELLETCVVNIALPNRSRRSKVSAR KIGLQKLWEWCIGIVQMTSLPWRVVIGRLGR LGHGELKDWSILLGECSLQTISKKLKDVCRM CGISAADSPSILSACLVAMEPQGSFVVMPDAV TMGSVFGRSTALNMQSSQLNTPQDASCTHIL VFPTSSTIQVAPANYPNEDGFSPNNDDMFVDL PFPDDMDNDIGILMTGNLHSSPNSSPVPSPGSP SGIGVGSHFQHSRSQGERLLSREAPEELKQQP LALGYFVSTAKAENLPQWFWSSCPQAQNQC PLFLKASLHHHISVAQTDELLPARNSQRVPHP
824	2174	A	6732	2440	365	LDSKTTSDVLRFVLEQYNALSWLTCNPATQD RTSCLPVHFVVLTQLYNAIMNIL VEEGLGRRRTPPGGRRGPVTPARPGPDSVRR
						RLLPPSSAAAFSSHRHNLLCSRRRGGGGGGGGGGGGGGGGGTIKRPGITGPTAATSPSGEPGNAASAP LSLLSPFPGQTTYQHPGVAEPSAYGGRDVAC ASLVFGRLQHRGGDRKRGLLGRSSGDAASD QPFRCRSGSTAGRLVKQMDFTEAYADTCSTV GLAAREGNVKVLRKLLKKGRSVDVADNRG WMPHEAAYHNSVECLQMLINADSSENYIKM KTFEGFCALHLAASQGHWKIVQILLEAGADP NATTLEETTPLFLAVENGQIDVLRLLLQHGAN VNGSHSMCGWNSLHQASFQENAEIIKLLLRK GANKECQDDFGITPLFVAAQYGKLESLSILIS SGANVNCQALDKATPLFIAAQEGHTKCVELL LSSGADPDLYCNEDSWQLPHAAAQMGHTKI LDLLIPLTNRACDTGLNKVSPVYSAVFGGHE DCLEILLRNGYSPDAQACLVFGFSSPVCMAFQ KDCEFFGIVNILLKYGAQNELHLAYCLKYEK FSIFRYFLRKGCSLGPWNHIYEFVNHAIKAQA KYKEWLPHLLVAGFDPLILLCNSWIDSVSIDT LIFTLEFTNWKTLAPAVERMLSARASNAWIL QQHIATVPSLTHLCRLEIRSSLKSERLRSDSYIS

SEQ ID NO: of nucl- eotide scq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QLPLPRSLHNYLLYEDVLRMYEVPELAAIQD G RIMGLFDRGVQMLLTTVGAFAAFSLMTIAVG TDYWLYSRGVCKTKSVSENETSKKNEEVMT HSGLWRTCCLEGNFKGLCKQIDHFPEDADYE
						ADTAEYFLRAVRASSIFPILSVILLFMGGLCIA ASEFYKTRIHNIII.SAGIFFVSAGLSNIIGIIVYIS ANAGDPSKSDSKKNSYSYGWSFYFGALSFIIA EMVGVLAVHMFIDRHKQLRATARAYTDYLQ ASAITRIPSYRYRYQRRSRSSSRSTEPSHSRDA SPVGIKGFNTLPSTEISMYTLSRDPLKAATTPT ATYNSDRDNSFLQVHNCIQKENKDSLHSNTA NRRTTPV
826	-	A	6744	3	5177	SDDLRTGLFQDVQDAESLKLPGVYEVLFYNE TEDCPGMMLWRYPEPRGLTLVRITPVPFNTT EDPDISTADLGDVLQDPCSLEYWDELQKVFV AFREFNLSESKVCELQLPDINLVNDQKKLVSS DLWRIVLNSSQNGADDQSSASESGSQSTCDPL VTPTALAACTRVDSCFTPWFVPSLCVSFQFAH LEFHLCHHLDQLGTAAPQYLQPFVSDRNMPS ELEYMIVSFREPHMYLRQWNNGSVCQEIQFL AQADCKLLECRNVTMQSVVKPFSIFGQMAVS SDVVEKLLDCTVIVDSVFVNLGQHVVHSLNT AIQAWQQNKCPEVEELVFSHFVICNDTQETL RFGQVDTDENILLASLHSHQYSWRSHKSPQL LHICIEGWGNWRWSEPFSVDHAGTFIRTIQYR GRTASLIIKVQQLNGVQKQIIICGRQIICSYLSQ SIELKVVQHYIGQDGQAVVREHFDCLTAKQK LPSYILENNELTELCVKAKGDEDWSRDVCLE SKAPEYSIVIQVPSSNSSIIYVWCTVLTLEPNS QVQQRMIVFSPLFIMRSHLPDPIIHLEKRSLGL SETQIIPGKGQEKPLQNIEPDLVHHLTFQAREE YDPSDCAVPISTSLIKQIATKVHPGGTVNQILD EFYGPEKSLQPIWPYNKKDSDRNEQLSQWDS PMRVKLSIWKPYVRTLLIELLPWALLINESKW DLWLFBGGKIVLQVPAGKIIIPPNFQEAFQIGIY WANTNTVHKSVAIKLVHNLTSPKWKDGGNG EVVTLDEEAFVDTEIRLGAFPGHQKLCQFCIS SMVQQGIQIIQIEDKTTIINNTPYQIFYKPQLSV CNPHSGKEYFRVPDSATFSICPGGEQPAMKSS SLPCWDLMPDISQSVLDASLLQKQIMLGFSPA PGADSSQCWSLPAIVRPEFPRQSVAVPLGNFR ENGFCTRAIVLTYQEHLGVTYLTLSEDPSPRV IHNRCPVKMLIKENIKDIPKFEVYCKKIPSECS IHHELYHQISSYPDCKTKDLLPSLLLRVEPLDE VTTEWSDAIDINSQGTQVVFLTGFGYVYVDV VHQCGTVFTTVAPEGKAGPILTNTNRAPEKIV TF/KMPITQLSLAVFDDLTHHKASAELLRILL DNIFLCVAPGAGPLPGEEPVAALFELYCVEIC CGDLQLDNQLYNKSNFHFAVLVCQGEKAEPI QCSKMQSLLISNKELEEYKEKCFIKLCITLNEG KSILCDINEFSFELKPARLYVEDTFVYYIKTLF DTYLPNSRLAGHSTHLSGGKQVLPMQVTQH ARALVNPVKLRKLVIQPVNLLVSIHASLKLVI ASDHTPLSFSVFERGPIFTTARQLVHALAMHY AAGALFRAGWVVGSLDILGSPASLVRSIGNG VADFFRLPYEGLTRGPGAFVSGVSRGTTSFVK HISKGTLTSITNLATSLARNMDRLSLDEEHYN RQEEWRRQLPESLGEGLRQGLSRLGISLLGAI AGIVDQPMQNFQKTSEAQASAGHKAKGVISG

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VGKGIMGVFTKPIGGAAELVSQTGYGILHGA GLSQLPKQRHQPSD\VHADQAPNSHVKYVW KMLQSLGRPEVHMALDVULVRGSQEHEGC LLLTSEVLFVVSVSEDTQQQAFPVTEIDCAQD SKQNNLLTVQLKQPRVACDVEVDGVRERLSE QQYNRLVDYITKTSCHLAPSCSSMQIPCPVVA AEPPPSTVKTYHYLVDPHFAQVFLSKFTMVK
827	2177	A	6748	2	1662	NKALRKGFP FVGAPRRGNPFGSPGNPGRHQGPCHRPRGTK ASGVSPTLWRPQAAATGLEMPSSGRALLDSP LDSGSLTSLDSSVFCSEGEGEPLALGDCFTVN VGGSRFVLSQQALSCFPHTRLGKLAVVVASY RRPGALAAVPSPLELCDDANPVDNEYFFDRS SQAFRYVLHYYRTGRLHVMEQLCALSFLQEI QYWGIDELSIDSCCRDRYFRRKELSETLDFKK DTEDQESQHESEQDFSQGPCPTVRQKLWNIL EKPGSSTAARIFGVISIIFVGVSIINMALMSAEL SWLDLQLLEILEYVCISWFTGEFVLRFLCVRD RCRFLRKVPNIIDLLAILPFYITLLVESLSG\SQT TQELVENVGAHCPGCLRLLRALRMLKAWGR HSTGLRSLGMTITQCYEEVGLLLLFLSVG\SIF STVEYFAEQSIPDTTFTSVPCAWWWATTSMT TVGYGDIRPDTTTGKIVAFMCILSGILVLALPI AINDRFSACYFTLKLKEAAVRQREALKKLTK NIATDSYISVNLRDVYARSIMEMLRLKGRER ASTRSSGGDDFWF
828	2178	A	6786	5672	1360	GTHPASSGPVPLPPAAVSAATREELGEPVFV TASSGFQSMHSSNPKVRSSPSGNTQSSPKSKQ EVMVRPPTVMSPSGNPQLDSKFSNQGKQGGS ASQSQPSPCDSKSGGHTPKALPGPGGSMGLK NGAGNGAKGKGKRERSISADSFDQRDPGTPN DDSDIKECNSADHIKSQDSQHTPHSMTPSNAT APRSSTPPHGQTTATEPTPAQKTPAKVVYVFS TEMANKAAEAVLKGQVEITVSFHIQNISNNK TERSTAPLNTQISALRNDPKPLPQQPPAPANQ DQNSSQNTRLQPTPPIPAPAPKPAAPPRPLDRE SPGVENKLIPSVGSSPASSTPLPPDGTGPNSTPN NRAVTPVSQGSNSSSADPKAPPPPPVSSGEPPT LGENPDGLSQEQLEHRERSLQTLRDIQRMLFP DEKEPTGAQSGGPQQNPGVLDGPQKKPEGPI QAMMAQSQSLGKGPGPRTDVGAPFGPGGHR DVPFSPDEMVPPSMNSQSGTIGPDHLDHMTP EQIAWLKLQQEFYEEKRRKPEQVVVQQCSLQ DMMVHQHGPRGVVRGPPPYQMTPSEGWAP GGTEPFSDGINMPHSLPPRGMAPHPNMPGSQ MRLPGFAGMINSEMEGPNVPNPASRPGLSGV SWPDDVPKIPDGRNFPPGQGIFSGPGRGERFP NPQGLSEEMFQQLAEKQLGLPPGMAMEGIR PSMEMNRMIPGSQRHMEPGNNPIFPRIPVEGP LSPSRGDFPKGIPPQMGPGRELEFGMVPSGM KGDVNLNVNMGSNSQMIPQKMREAGAGPEE MLKLRPGGSDMLPAQQKMVPLPFGEHPQQE YGMGPRPFLPMSQGPGSNSGLRNLREPIGPDQ RTNSRLSHMPPLPLNPSSNPTSLNTAPPVQRG LGRKPLDISVAGSQVHSPGINPLKSPTMHQVQ SPMLGSPSGNLKSPQTPSQLAGMLAGPAAAA SIKSPPVLGSAAASPVHLKSPSLPAPSPGWTSS PEPPLQSPGIPPNHKAPLTMASPAMLGNVESG GPPPTASQPASVNIPGSLPSSTPYTMPPEPTL SQNPLSIMMSRWSKFAMPSISNPGYNHDAI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion KTVASSDDDSPPARSPNLPSMNMPGMGINT QNPRISGPNPVVPMPTLSPMGMTQPLSHSNQ MPSPNAVGPNIPPHGVPMGPGLMSHNPIMGH GSQEPPMVPQGRMGFPQGFPPVQSPPQVPFP HNGPSGQGSFPGGMGFPGGPLGRPSNLPQ SSADAALCKPGGPGGPDSFTVLGNSMPSVFT DPDLQEVIRPGATGIPEFDLSRIIPSEKPSQTLQ YFPRGEVPGRKQPQGPGFSHMQGMMGEQ APRMGLALPGMGGPGPVGTPDIPLGTAPSMP GHNPMRPPAFLQQGMMGPHIRMMSPAQST MPGQPTLMSNPAAAVGMIPGKDRGPAGLYT HPGPVGSPGMMMSNAQGMMGPNNRTS
829	2179	A	6797	433	3	ASFFNFSICICKIILEVGPPVGHPAHDDVGGRH GPGGR/GSRSPRSLQCAPGGGRRSGCPAGSSP ASTCPPSPGGSGADRFGPSPPPPSREAAPTAG AAASSTSSGASCPPVPASSRWGVRSRTRSGSG
830	2180	A	6800	3	1911	GEREPRDRPSERPRLV LPERAFGPRTPRAPRRRRRRLLLSPPPRPPPPL DREPRAPGPWLCPSRAGTAQDPARIRERRGR VAGGAAGPAMELRARGWWLLCAAAALVAC ARGDPASKSRSCGEVRQIYGAKGFSSSIDVPQ AEISGEHLRICPQGYTCCTSEMEENLANRSHA ELETALRDSSRVLQAMLATQLRSFDDHFQHL LNDSERTLQATFPGAFGELYTQNARAFRDLY SELRLYYRGANLHLEETLAEFWARLLERLFK QLHPQLLLPDDYLDCLGKQAEALRPFGEAP\ RELRLRATRA\FVAAR\SFVQGLGVAS\DVVR KVAQVPLG\PECSRAVIEAGSYC\ALHCVGVP GARPCPDYCRNVLKGCLANQADLDAEWRNL LDSMVLITDKFWGTSGVESVIGSVHTWLAEA INALQDNRDTLTAKVIQGCGNPKVNPQGPGP EEKRRGKLAPRERPPSGTLEKLVSEAKAQL RDVQDFWISLPGTLCSEKMALSTASDDRCWN GMARGRYLPEVMGDGLANQINNPEVEVDIT KPDMTIRQQIMQLKIMTNRLRSAYNGNDVDF QDASDDGSGSGSGDGCLDDLCGRKVSRKSSS SRTPLTHALPGLSEQEGQKTSAASCPQPPTFL LPLLLFLALTVARPRWR
831	2181	A	6808	2	1522	ASRHGMTPGALLMLLGALGPPLAPGVRGSEA EGRLREKLFSGYDSSVRPAREVGDRVRVSVG LILAQLISLNEKDEEMSTKVYLDLEWTDYRLS WDPAEHDGIDSLRITAESVWLPDVVLLNNND GNFDVALDISVVVSSDGSVRWQPPGIYRSSCS IQVTYFPFDWQNCTMVFSSYSYDSSEVSLQT GLGPDGQGHQEIHIHEGTFIENGQWENIHKPS RLIQPPGDPRGGREGQRQEVIFYLIIRRKPLFY LVNVIAPCILITLLAIFVFYLPPDAGEKMGLSIF ALLTLTVFLLLLADKVPEISLSVPIIIKYLMFT MVLVTFSVILSVVVLNLHHRSPHTHQMPLWV RQIFIHKLPLYLRLKRPKPERDLMPEPPPHCSSP GSGWGRGTDEYFIRKPPSDFLFPKPNRFQPEL SAPDLRRFIDGPNRAVALLPELREVVSSISYIA RQLQEQEDHDALKEDWGFVAMVVDRLFLW TFIIFTSVGTL\VIFLDATYHLPPPDPFF
832	2182	A	6824	71	1079	ETMAKNPPENCEDCHILNAEAFKSKKICKSLK ICGLVFGILALTLIVLFWGSKHFWPEVPKKAY DMEHTFYSNGEKKKIYMEIDPVTRTEIFRSGN GTDETLEVHDFKNGYTGIYFVGLQKCFIKTQI KVIPEFSEPEEEIDENEEITTTFFEQSVIWVPAE KPIENRDFLKNSKILEICDNVTMYWINPTL\IS

SEQ ID NO: of NO: of nucleotide peptide cotide sequence uenc	GEDLHFP HARQAS CIYCRR GRVICRV ETARIGP KDEKE RRFRVR MERIGE PHHDHII
nucleotide eotide sequence In USSN 09/496 1914 1914	GEDLHFP HARQAS CIYCRR GRVICRV ETARIGP EKDEKE RRFRVR IMERIGE PHHDIIII
eotide sequence	GEDLHFP HARQAS CIYCRR GRVICRV ETARIGP EKDEKE RRFRVR IMERIGE PHHDIIII
sequence Sequence 194	GEDLHFP HARQAS CIYCRR GRVICRV ETARIGP KDEKE RRFRVR MERIGE PHHDHII LGGRSR
uence 914 ng to first amino acid residue of peptide residue of peptide sequence Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon /=possible nucleotide deletion, \=possible nucleotide insertion GTFAKQLHHNFAFIILVSELQDFEEEA ANEKKGIEQNEQWVVPQVKVEKTRI EEELPINDYTENGIEFDPMLDERGYC GNRYCRRVCEPLLGYYPYPYCYQGC IMPCNWVARMLGRV EAEGEQVCGAKCCGDAPHVENREEIA GVMESKEERALNNLIVENVNQENDE QVANKGEPLALPLNVSEYCVPRGNR QPILQYRWDIMHRLGEPQARMREEN EVRQLMEKLREKQLSHSLRAVSTDPDEFCLMP PNGVALHLPGAAVIPNTNYMFQDA GSREESPAPSRAPASASLWRRLVVVF HAAAAAQAAAAQAAHAEAADSWY AEHFRTSSPPKIRLCVHCLQAVFPFKIRTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAJGSQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMILLLMERKLQEVHPLLTLCGGIVE PIQKESLRVFFLVLQVTHYLDAGQVK	GEDLHFP HARQAS CIYCRR GRVICRV ETARIGP KDEKE RRFRVR MERIGE PHHDHII LGGRSR
amino acid residue of peptide sequence	GEDLHFP HARQAS CIYCRR GRVICRV ETARIGP KODEKE RRFRVR MERIGE PHHDHIII
residue of peptide sequence	GEDLHFP HARQAS CIYCRR GRVICRV ETARIGP KODEKE RRFRVR MERIGE PHHDHIII
peptide peptide possible nucleotide deletion, -possible nucleotide insertion GTFAKQLHHNFAFIILVSELQDFEER ANEKKGIEQNEQWVVPQVKVEKTRI EEELPINDYTENGIEFDPMLDERGYC GNRYCRVCEPLLGYYPYPYCYQGC IMPCNWWVARMLGRV GNRYCRVCEPLLGYYPYPYCYQGC IMPCNWWVARMLGRV EAEGEQVCOAKCCGDAPHVENREEI GVMESKEERALNNLIVENVNQENDE QVANKGEPLALPLNVSEYCVPRGNR QPILQYRWDIMHRLGEPQARMREEN EVRQLMEKLREKQLSHSLRAVSTDP DEFCUMP DEFCUMP DEFCUMP AAAAAQAAAAQAAHAEAADSWYAEHFRTSSPPKIRLCVHCLQAVFPKI RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSLLYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	GEDLHFP HARQAS CIYCRR GRVICRV ETARIGP KADEKE RRFRVR MERIGE PHHDHIII
833 2183 A 6846 116 602 EAEGEQVCGAKCCGDAPHVENREEI GVANKGEPLALPLNVSEYCVPRGNR QPILQYRWDIMHRLGEPQARMREEN EVRQLMEKLREKQLSHSLRAVSTDP DEFCLMP 834 2184 A 6851 3 2024 PNGVALHLPGAAVIPNTNYMFQDA GSREESPAPSRAPASASLWRLVVV HAAAAAQAAAAQAAHAEAADSWY AEHFRTSSPPKIRLCVHCLQAVFPKI RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	GEDLHFP HARQAS CIYCRR GRVICRV ETARIGP EKDEKE RRFRVR IMERIGE PHHDIIII
833 2183 A 6846 116 602 EAEGQVCGAKCCGDAPHVENREEI GVANKGEPLALPLNVSEYCVPRGNR QVANKGEPLALPLNVSEYCVPRGNR QVANKGEPLALPLNVSEYCVPRGNR QVILQYRWDIMHRLGEPQARMREEN EVRQLMEKLREKQLSHSLRAVSTDP DEFCLMP 834 2184 A 6851 3 2024 PNGVALLHLPGAAVIPNTNYMFQDA GSREESPAPSRAPASASLWRRLVVVI HAAAAAQAAAAQAAHAEAADSWY AEHFRTSSPPKIRLCVHCLQAVFPKI RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQU LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	HARQAS CIYCRR BRVICRV ETARIGP EKDEKE RRFRVR IMERIGE PHHDHIII LGGRSR
833 2183 A 6846 116 602 EAGGEQVCGAKCCGDAPHVENREEJ GVMESKEERALNNLIVENVNQENDE QVANKGEPLALPLNVSEYCVPRGNR QPILQYRWDIMHRLGEPQARMREEN EVRQLMEKLREKQLSHSLRAVSTDP DEFCLMP PNGVALLHLPGAAVIPNTNYMFQDA GSREESPAPSRAPASASLWRRLVVVI HAAAAAQAAAAQAAHAEAADSWY AEHFRTSSPPKIRLCVHCLQAVFPFKJ RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEJ PIQKESLRVFFLVLQVTHYLDAGQVK	HARQAS CIYCRR BRVICRV ETARIGP EKDEKE RRFRVR IMERIGE PHHDHIII LGGRSR
833 2183 A 6846 116 602 EAEGEQVCGAKCCGDAPHVENREEI GVMESKEERALNNLIVENVNQENDE QVANKGEPLALPLNVSEYCVPRGNR QPILQYRWDIMHRLGEPQARMREEN EVRQLMEKLREKQLSHSLRAVSTDP DEFCLMP 834 2184 A 6851 3 2024 PNGVALLHLPGAAVIPNTNYMFQDA GSREESPAPSRAPASASLWRRLVVV HAAAAAQAAAAQAAHAEAADSWY AEHFRTSSPPKIRLCVHCLQAVFPFXI RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	CIYCRR BRVICRV ETARIGP EKDEKE RRFRVR IMERIGE PHHDHIII
833 2183 A 6846 116 602 EAEGEQVCGAKCCGDAPHVENREER GVMESKEERALNNLIVENVNQENDE QVANKGEPLALPLNVSEYCVPRGNR QPILQYRWDIMHRLGEPQARMREEN EVRQLMEKLREKQLSHSLRAVSTDP DEFCLMP 834 2184 A 6851 3 2024 PNGVALLHLPGAAVIPNTNYMFQDA GSREESPAPSRAPASASLWRRLVVV HAAAAAQAAAAQAAHAEAADSWY AEHFRTSSPPKIRLCVHCLQAVFPKX RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQU LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	ETARIGP EKDEKE EKRFRVR IMERIGE PHHDHIII
833 2183 A 6846 116 602 EAEGEQVCGAKCCGDAPHVENREER GVMESKERALNNLIVENVNQENDE QVANKGEPLALPLNVSEYCVPRGNR QPILQYRWDIMHRLGEPQARMREEN EVRQLMEKLREKQLSHSLRAVSTDP DEFCLMP 834 2184 A 6851 3 2024 PNGVALLHLPGAAVIPNTNYMFQDA GSREESPAPSRAPASASLWRRLVVVI HAAAAAQAAAAQAAHAEAADSWY AEHFRTSSPPKIRLCVHCLQAVFPFKI RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	ETARIGP EKDEKE RRFRVR IMERIGE PHHDHIII
833 2183 A 6846 116 602 EAEGEQVCGAKCCGDAPHVENREEI GVMESKERALNNLIVENVNQENDE QVANKGEPLALPLNVSEYCVPRGNR QPILQYRWDIMHRLGEPQARMREEN EVRQLMEKLREKQLSHSLRAVSTDP DEFCLMP 834 2184 A 6851 3 2024 PNGVALLHLPGAAVIPNTNYMFQDA GSREESPAPSRAPASASLWRLVVVI HAAAAAQAAAAQAAHAEAADSWY AEHFRTSSPPKIRLCVHCLQAVFPFKI RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSLLYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	KDEKE RRFRVR IMERIGE PHHDHIII LGGRSR
GVMESKEERALNNLIVENVNQENDE QVANKGEPLALPLNVSEYCVPRGNR QPILQYRWDIMHRLGEPQARMREEN EVRQLMEKLREKQLSHSLRAVSTDPI DEFCLMP 834 2184 A 6851 3 2024 PNGVALLHLPGAAVIPNTNYMFQDA GSREESPAPSRAPASASLWRLVVVI HAAAAAQAAAAQAAHAEAADSWYI AEHFRTSSPPKIRLCVHCLQAVFPFKI RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	KDEKE RRFRVR IMERIGE PHHDHII LGGRSR
834 2184 A 6851 3 2024 PNGVALLHLPGAAVIPNTNYMFQDA 834 CARRENT SEPPKIRLCVHCLQAVFPKIR 834 CARRENT SEPPKIRLCVHCLQAVFPKIR 834 CARRENT SEPPKIRLCVHCLQAVFPKIR 834 CARRENT SEPPKIRLCVHCLQAVFPKIR 834 CARRENT SEPPKIRLCVHCLQAVFPKIR 834 CARRENT SEPPKIRLCVHCLQAVFPKIR 834 CARRENT SEPPKIRLCVHCLQAVFPKIR 834 CARRENT SEPPKIRLCVHCLQAVFPKIR 834 CARRENT SEPPKIRLCVHCLQAVFPKIR 834 CARRENT SEPPKIRLCVHCLQAVFPKIR 834 CARRENT SEPPKIRLCVHCLQAVFPKIR 835 CARRENT SEPPKIRLCVHCLQAVFPKIR 836 CARRENT SEPRENT SEPPKIRLCVHCLQAVFPKIR 837 CARRENT SEPPKIRLCVHCLQAVFPKIR 837 CARRENT SEPPKIRLCVHCLQAVFPKIR 836 CARRENT SEPPKIRLCVHCLQAVFPKIR 837 CARRENT SEPPKIRLCVHCLQAVFPKIR 837 CARRENT SEPPKIRLCVHCLQAVFPKIR 837 CARRENT SEPPKIRLCVHCLQAVFPKIR 836 CARRENT SEPPKIRLCVHCLQAVFPKIR 837 CARRENT SEPPKIRLCVHCLQAVFPKIR 837 CARRENT SEPPKIRLCVHCLQAVFPKIR 837 CARRENT SEPPKIRLCVHCLQAVFPKIR 838 CARRENT SEPPKIRLCVHCLQAVFPKIR 848 CARRENT SEPPKIRLCVHCLQAVFPKIRL 849 CARRENT SEPPKIRLCVHCLQAVFPKIRL 849 CARRENT SEPPKIRLCVHCLQAVFPKIRL 849 CARRENT SEPPKIRLCVHCLQAVFPKIRL 849 CARRENT SEPPKIRLCVHCLQAVFPKIRL 840 CARRENT SEPPKIRLCVHCLQAVFPKIRL 840 CARRENT SEPPKIRLCVHCLQAVFPKIRL 840 CARRENT SEPPKIRLCVHCLQAVFPKIRL 840 CARRENT SEPPKIRLCVHCLQAVFPKIRL 840 CARRENT SEPPKIRLCVHCLQAVFPKIRL 840 CARRENT SEPPKIRLCVHCLQAVFPKIRL 840 CARRENT SEPPKIRLCVHCLQAVFRAV SEPPKIRL 840 CARRENT SEPPKIRLCVHCLQAVFRAV SEPPKIRL 840 CARRENT SEPPKIRLCVHCLQAVFRAV SEPPKIRLCVHCLQAVFRAV SEPPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHCLQAVFRAV SEPKIRLCVHC	RRFRVR IMERIGE PHHDHIII LGGRSR
QPILQYRWDIMHRLGEPQARMREEN EVRQLMEKLREKQLSHSLRAVSTDPI DEFCUMP 834 2184 A 6851 3 2024 PNGVALLHLPGAAVIPNTNYMFQDA GSREESPAPSRAPASASLWRRLVVVI HAAAAAQAAAAQAAHAEAADSWY AEHFRTSSPPKIRLCVHCLQAVFPKI RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	MERIGE PHHDIIII LGGRSR
EVRQLMEKLREKQLSHSLRAVSTDPI DEFCUMP 834 2184 A 6851 3 2024 PNGVALLHLPGAAVIPNTNYMFQDA GSREESPAPSRAPASASLWRRLVVVI HAAAAAQAAAAQAAHAEAADSWY AEHFRTSSPPKIRLCVHCLQAVFPKI RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	PHHDHIII LGGRSR
834 2184 A 6851 3 2024 PNGVALLHLPGAAVIPNTNYMFQDA GSREESPAPSRAPASASLWRRLVVVI HAAAAAQAAAAQAAHAEAADSWY AEHFRTSSPPKIRLCVHCLQAVFPKI RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	LGGRSR
834 2184 A 6851 3 2024 PNGVALLHLPGAAVIPNTNYMFQDA GSREESPAPSRAPASASLWRRLVVVI HAAAAAQAAAAQAAHEAADSWY AEHFRTSSPPKIRLCVHCLQAVFPFKI RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	
GSREESPAPSRAPASASLWRRLVVVI HAAAAAQAAAAQAAHAEAADSWY AEHFRTSSPPKIRLCVHCLQAVFPFKI RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	
HAAAAQAAAAQAAHAEAADSWY AEHFRTSSPPKIRLCVHCLQAVFPFKJ RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEJ PIQKESLRVFFLVLQVTHYLDAGQVK	
AEHFRTSSPPKIRLCVHCLQAVFPFKI RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	
RTHLQLGSVLYHHTKNSEQARSHLE QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	
QQIPQFEDVKFEAASLLSELYCQENS LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVEI PIQKESLRVFFLVLQVTHYLDAGQVK	KAWLIS
LLRKAIQISQQTPYWHCRLLFQLAQL LVSACDLLGVGAEYARVVGSEYTRA GMLLLMERKLQEVHPLLTLCGQIVE PIQKESLRVFFLVLQVTHYLDAGQVK	
GMLLLMERKLQEVHPLLTLCGQIVE PIQKESLRVFFLVLQVTHYLDAGQVK	
PIQKESLRVFFLVLQVTHYLDAGQVK	LFLLSK
	NWQGN
LKQLQQCIQTISTLHDDEILPSNPADL	
KEHMCVLVYLVTVMHSMQAGYLEK	AOKYT
DKALMQLEKLKMLDCSPILSSFQVIL	LEHIIM
CRLVTGHKATALQEISQVCQLCQQSI	PRLFSN
HAAQLHTLLGLYCVSVNCMDNAEA	QFTTAL
RLTNHQELWAFIVTNLASVYIREGNR	HQEVV\
LYSLLERINPDHSFPVSSHCLRAAAF)	/VRGLF
SFFQGRYNEAKRFLRETLKMSNAEDI	LNRLTA
CSLVLLGHIFYVLGNHRESNNMVVPA	AMQLAS
KIPDMSVQLWSSALLRDLNKACGNA	
AAQMHQNFSQQLLQDHIEACSLPEH	NLITWT
DGPPPVQFQAQNGPNTSLASLL	
835 2185 A 6855 334 1268 PTRRPILPLTSPKAISVPSPLQGKQHTI	VKSCL
SVSGIGGFLVSLSSRMKLQTLAVSVT	
AYVPCQTQDRDALRLTLEQIDLIRRM	
LELVTSAKALNDTQKLACLIGVEGGI	
LSILRTFYMLGVRYLTLTHTCNTPWA	
GVHSFYNNISGLTDFGEKVVAEMNR	
DLSHVSDAVARRALEVSQAPVIFSHS	
CNSARNVPDDILQLLEEERWAFVMV	
LIQWQPIRPMCSTVADHFDHIKAVIG	
have the common the property and an include the common transfer of t	
RPEIRLLPLPMTGACGAVAASRTGSS	
PNGHGGKGSGLANGLAGNP\GHLGLo	J35FGT
	TTO TIME
- I I I I I I I I I I I I I I I I I I I	
APFPGLVQRRSRLLIVSQVRYFLKNK	
NEDGLTALHQCCIDNFEEIVKLLLSH(
AKDNELWTPLHAAATCGHINLVKILV	
DLLAVNSDGNMPYDLCEDEPTLDVIE	
QGITQEKINEMRVAPEQQMIADIHCM	
DLDWIDAQGATLLHIAGANGYLRAA	
GVRVDVKDWDGWEPLHAAAFWGQI	
LLVSHGANLNARTSMDEMPIDLCEE	
LLELK\HKHDVIMKSQLRHKSSLSRR	SHROA
S/SVGKVVRRTQPVGTGPNL\YRKEYI	
LWQRSA\AEDQRTSTYNGDIRET\RTL	E/GEEAI

SRQ ID No. of N							(A. Alesian Ce-Custing							
NO. of mole- onder graphic counter graphic counter graphic counter graphic counter graphic counter graphic gra	CEO ID	SEO ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A-Alanine C-Cysteine,							
mod- outlide seq- uence 11			hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,							
USSN Og496 914 1058 09496 914 1058 09496 914 1058 09496 914 1058 09496 1059 1059 1059 1059 1059 1059 1059 1059						location	F=Phenylalanine, G=Glycine, H=Histidine,							
### Denice 1944 194		1 * -	1		**	corresponding	I=Isoleucine, K=Lysine, L=Leucine,							
teine to the common and the residue of peptide sequence peptides and properties of peptide sequence peptides sequence pe			Į		•		M=Methionine, N=Asparagine, P=Proline,							
mino acid residue of peptide sequence s		uence		1	•		O=Glutamine, R=Arginine, S=Serine,							
residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence sequen	uence	1	1	314			T=Threonine, V=Valine, W=Tryptophan,							
### ### ##############################			{	ŀ			V=Tyrosine X=IInknown *=Stop codon.							
nucleotide insertion PREFILE/ENVILLSEPT/KIPRGELDMPVENGLR APVSAYQYALANGDVWKVIE/PDYSMAYG NPVADATPWSSYKGSPOTILLE/KRORAA AKILSHPFLSTHIJGSSMARTGESSSEGKAPLI GERTSPYSSNGTSVYLYTSGDPPLIKKAPI EDMEEKVHOCCNIS 838 2188 A 6865 6291 739 AGPLERWOGAMALQLWALTLICILJGAGAS LRPRILDFRSKEINHLAVDEASGVVYLGA VNALYQLDAKIQLEQQVATGPVLDMKKCTP PEASQCHEAEMTDVYNQLLLVDPPRKRI,VE CGQLKGRCALRALSNIS.RLFTEDGSGEKSF VASNDEGVATGVLSTOFGGRVLFVGKG NGPHDMGIIVSTRLDRTBREAFEAYTDHAT YKAGYLSTNTOQFVAAFBOGTVJFVYNQCU KHARNRTLLARMCREDRYYSYLEMDLQC RDPHHAAAFGTCLAASVAAPGSGRVLYVGK RDPHHAAAFGTCLAASVAAPGSGRVLYVAF SRDSRSSGGPGAGLCLFPLDEVHAKMEANEN ACYTGTREARDIFYEPHGOPLOGHAPOSK SFPCGSEEILPYPLGSRDGLRGTAVLQRGGLN LTAVTVAADNNHTVALGTSDGRLKVYLTP DOTSSEYDSILVENKRVKRDLVLSGDLGSLY AMTODKVRILPVQELSTYPTTOCROSSODPY CGWCVVGRGTTRKAECFRAEEASHWLWSRS XSCVAVTSAQPOMNSRRAQGSVQLVTSPLAA LSEEDELLCLFGESPHPARVEGEAVICNSPSS PYTPROGDHVANTIQLLRRGNHETSYCYPF YDCRQAMSLEENLFCISCVSNRWTCQWDLA HECREASNPEGDIVARHMEDSCCPGLGPS LVPMHETDVFGGRNLDTVKGSSLHVGISD LKFMERVTMQSSGTFARFTKLSIGHAMETL PLHLYVKSVGKNIDSKLHVTLYDCSFGRSDC SLCRAAANPDYRCAWGGGGRCVYBALCNTT SECPPVITTLIOPETOPLGGGRITLLGSNLGVQ AGDIGSVAGRNCGSPREYSSTRIVCVUEA AETPFTGGVFUDVFGKLGRSPPNOPTFQQP KRLSVEPQQDPAGGGGRCVYBALCNTT SECPPVITTLIOPETOPLGGGRITLLGSNLGVQ AGDIGSVAGRNCGSPREYSSTRIVCVUEA AETPFTGGVFUDVFGKLGRSPPNOPTFQQP KRLSVEPQQDPAGGGRSCVYBALCNTT SCCPPVITTLIOPETTYREDPVLAFE LSWGPRAESSQPRIVVGTDYFINDTK VFLSFAVDEEPEAYNTVALFIGHDDHALLRT LPLHLYVKSVGKNDSKLHVTLYDCSFGRSDC SLCRAAANPDYRCAWGGGRSCVYBALCNTT SCCPPVITTLIOPETTYREDPVLAFE PLRSFASGGRSINVTQGTSLIJGRFAMVVIAEP LSWGPRAESSQPRIVVGTDYFINDTK VFLSFAVDEEPEAYNTVALFIGHDDHALLRT LEGAGEVPUPPTFRENTGGOKKQVNKLIRAR GREWVYLDVKSVYCKSRSOKQNKRUTLAR GREWVYLDVKSVYCKSRSOKGNKYNILLWRSTVV ERMLSNWMSICLYQYKKSRDCHVKLIFRAI LEGNGVKRYNTLMHTYNVE VVXIANSVYCYWKSRQQAERYEKIKSQLEG LESSVROPCKREFTDIAMEMDGOTINVHEAG IPVLDVKTYTTDRVFFLPSKOGDKDVMITGKL DPPERRYVRQALYQYKSNYLMKRRSTVV ERMLSNWMSICLYQYKKSRGCHVKKIRSETIVV ERMLSNWMSICLYQYKKSRGCOKRPENSKLINFHIT LLENGVYKHQYGGORGCORPADSVLJEWRG STAQLISDLDLTSGRGRWKKNVILLMHTYNFO VXXIANSTVARSTVARSTVALFERVART LEENRYWHLYRPTDEVDEGKSKGSVKKE		1	Ì	ì		sequence	/							
PIPRIEKEVILSEPTIKIPRGELDMFVENOLA APVSAYOYALANDOWKVIEPYDYSMAYO NPGVADATPPWSSYKEQSPOTILELKRQRAA AKLISHIPISTHILGSMAKTOESSEGKAPLI GGRTSPYSSNGTSVYTYTVISGDPPLIKFKAPI EAMEEVHOCKIS 338 2188 A 6865 6291 739 AGPLERVISGAMALQLWALITLICHLGAGAS LURRIKLIDEFSEKELIHALAVDEASGVVYLGA VANLYQLDAKLIQLEQQVATGPVLDNKKCTP PIEASCHEAEMTDAVIQLLIVDPRKRILVE CGQLLKGINCALRALSNISLRIFYEDGSGEKSF VASNDEGVATVGLVSSTOPGGDRVLTYCKG NGPHDMGIVSTRLIDRIDSREAFEAYTDHAT YKAGYLSTNTQGFVAAFEDGFVYFYLYGKG NGPHDMGIVSTRLIDRIDSREAFEAYTDHAT YKAGYLSTNTQGFVAAFEDGFVYTYNQD KHARNRILARNCEDPHYYSYLEMDLQC RDPDHAAAGTCLAASVAAPGGGRVLYAVF SRISKSSGGFAGALLFPLIEPVHAKMEANEN ACYTGTREARDIFYKPHEDDIOCGGHAPGSK SFCOSEEH PYBLGSRDGLKGTAVLQRGGIAN LTAYTVAAENHTVALGTSDGRILKYLTP DGTSSSFYDSILVENIKKYRROLVLSGDLGLY AMTQDKVFELPVQECLSYPTCTQCNISQDFY CGWCVVSGRCTKAEGCRAEASAFWL VSRS KSCVAVTSAQRQNMSRRAQGBVQLTVSFLDR KSCVAVTSAQRQNMSRRAQGBVQLTVSFLDR LSEEDELLCFGESPFPHARVEGGAVCNSPSS IPVTPPGQDHVAVTIQLLLRRGNIFLTSYQYFF YDCRQAMSLEENHLEGSVSNNWTCONSPSS IPVTPPGQDHVAVTQLLLRRGNIFLTSYQYFF YDCRQAMSLEENHLEGSVSNNWTCONSPSS IPVTPPGQDHVAVTQLLLRRGNIFLTSYQYFF YDCRQAMSLEENHLEGSVSNNWTCONSPSS IPVTPRGDRIFTYRRYGGGAVCONDLR YHECREASPNEDGIVRAHMEDSCOPCLOPSY LIPMMHHEDVNPGGKLDVKSNBLINVICNIEA AETFFTGGGVEVDYRGKLDVTNTCSSLAVGD LLKFMEPVTMQESGTFARFTFRISHIDAMETL PHLYVKSYGKIDISKLIVVILTNDSGGRSDC SLCAANPDYRCAWGGGGSRCVFEALCNIT SCCPPTVTRQPFTGTGFGGGRTHLLGSNLGVQ AGDIQRISVAGRKCSFQPEXYSVSTRIVCVIEA AETFFTGGGVEVDYRGGQLGCVTGPQATRG QMLLEVSYGGSVPTRGGFTTYRENPVLRAFE PLRSFASGGSRSNIVTQGFSLNGVFTQQP KYLSVFPQQGPQAGGTTT.THGTHLDTGSQED VXVTLNOVPCKVTKFGQQLQCVTGPQATRG QMLLEVSYGGSVPTRGGFTYRENPVLRAFE LGSWCPPREASSLQPMTVVGTDYYFINDTK VYFLSPAVEPEPPAYNTVLFIEMDGHALLRT EAGAFEYYPDTTENFTTGGVEXKQVNKLIRAR GREWVYGNAKSPLAMRRSETVV ERMLSNWMSICLYQYLKDSRODROWMITGKL DPEPRFVVRQALYQSSNLNSKELNPHT LUENQFESABAKVYTASLITVALHGKLEYTY DMHTTLELLEGVYCKSPCSWPTSSVVLEWPG STAQILSDLDLTSQRGRWKKNTILMFTYNY VVVIANSVCYWBKSQQAERYEKKSQLG LESSYROPCSCRFFTDDAVKSPNLAMRSETVV ERMLSNWMSICLYQYLKSPRAGGRHALL LEENRYWHLVRPTDEVGBGRHALL EEDRIVKHUYRPTDEVGBGSKRGSKKGSWEKE		1	1											
APVSAYQYALANGDVWKVHEVPDYSMAYG NGVADATPWSSYREGSPOTILLEKRORAA AKILSHPFLSTHLOSSMARTGESSSEGKAPLI GRITSPYSSNGTSVYTYTSGDPPLIKRKAPI EEMEEKVIGCCRIS AGPLEPRYGGMALQLWALTLIGLIGAGAS LGPLEPRYGGMALQLWALTLIGLIGAGAS LGPLERPYGGMALQLWALTLIGLIGAGAS LARRILDFFSSKELNHLAVDEASGVVYLGA VNALYQLDAKUQLGQVATGPVLDNKKCTP PIEASGCHEAEMTDNYNQLLLVDPPRKRI.VP CGQLLKGRCALRAISHLKYPEDGSCEKSF VASNDEGVATGVULNSKLTPPHOSCEKSF VASNDEGVATVGUSTSTOPGGDRVLFVGKG NGPHDNGIVSTRLDRTDSRAFEAYTDHAT YKAGYISTNTQGVAAFDGDPVYFVFNQDQ KHARNKTLLARMCREDNYYSYLEMDLQC RDPDHAAAGTCLAAAFDGPYVFFVFNQDQ KHARNKTLLARMCREDNYYSYLEMDLQC SRDSSSSGGGAGLCLFPLDEVHAKMEANRN ACYTGTREARDIFYPHFOGQCGGRAVLARGASK SFCGSEEILPYPLGSRDGLRGTAVLGRGGLN LTATVAADANNTVALHGTSDGRILKVYLTP DGTSSEYDBILVENRKVKRDLVLSGDLGSLY AMTODKVFRLPVDGCSSYPTCTTQCRISODPY CGWCVVEGRCTRKAECFRAEEASHWLWSRS XSCVAVTSAQPONMSRRAQGSVQTVSPLAA LSEEDELLCLFGESPPHPARVEGAVICNSPSS PYTPROGDIVANTIQLLRRGNHLTSTYOPPF YDCRQAMSLEENLFCISCVSNRWTCQWDLR YHECRAESNPEGDIVATHAHEDSCCPLGFSS LLYFMHETDNYFQGKNLDTVKGSSLHVGISD LLKFMEVTVMGSSGTAFAFTRYSISHDAMELL PHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAAANPPYRCAWGGGSRCVYEALCNTT SECPPVITELOPSTOPLGGGRITLLGSNLGVQ ADDIGNSVAGRNCSPCREYSYSTTIVCVEA ASTEPTTGGVEVDVFGKLGRSPPNVQFTQQP RGLSSPRQQGPAGGGGSCVYCALCNTT SCCPPVITELOPSTOPLFTCRPPLIAFE QWCHYSTAGSTRAFTRYSHISHDAMELL LYHLYNSYGCSVPPNOPITTAGFTPHALTSTOPLAGE LLGSWGPREASSLQPMVVGTDYPHODIX VVFLNAAAPPYRCAWGTDYPTHOGS LLKFMERTART EAGAFEVYPDTFENTIGGVKRQVNLIRAR GNILKKAMTQGAEAFVQBETTMAKTLIFT DLYCEPPEVQPTFRENTGGVKRQVNLIRAR GNILKKAMTQGAEAFVQBETYSKRMVILLEN DLYCEPPEVQPTFRENTGGVKRQVNLIRAR GNILKKAMTQGAEAFVGAECTKMRSETVV GRSEWVLGNVENTYSNSVPSLLIPLTVVPM VVVIANSVYCYWRKSQQAEREYEKIKSQLEG LESSYRDRCKKEFTDIATWSDVHLEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DPEPERPVPQGALVGRPSNLNSSEN,INPHT LIEMQGFSAARKVYTASLITVALHGKLEVYT DMHTHI-ELLEQVYCARPYGSNLNSSEN,INPHT LIEMQGFSAARKVYTASLITVALHGKLEVYT DMHTHI-ELLEQVYCARPYGDTVLARPG STAQLISDLDLTSGRGRWKRNVILLMFYNFOSVL VKERIDOVYRGGCGFOROPDSODQLDFGERHALL EENRWHYH WPTDEVGEGSKRGSVKEKE EYALTEND THIL SKGTGLOGPDDSODDLFGERHALL EENRWHYH WPTDEVEGSKRGSVKEKE			_		sequence	<u></u>	BUCKOUGE INSCHOOL							
NPGVADATPPWSSYKEGSROTLLELKRQRAA AKLLSHPLSTHLGSMANTGESSEGKAPL GGRTSPYSSNGTSYYTYTOSGPPLLKFKAP EEMEEK-VHGCRIS GGRTSPYSSNGTSYYTYTOSGPPLLKFKAP EEMEEK-VHGCRIS GGRTSPYSSNGTSYYTYTOSGPPLLKFKAP EEMEEK-VHGCRIS AGPLERVGGAMALQLWALTLLGLIGAGAS LRFKLDFFSSKELJHALAVDEAGGVVYJGA VNALYQLDAKLQLEQQVATGPVLDNKKCTP PELASQCHEAENTDNYNQLLLVDPPRRILVP GCQLLKGLCALRALSNISLRLFYEDGSGEKSF VASNDEGVATVQLVSSTOPGGRVLFVGKG NGPHDKGINSTRLLDRIDSREAFRAYTDHAT YKAGYLSTNTOQFVAAFEDGFYVFFVROQD KHARNRTLLARNGEDRYNYSVLEMDLQC RDPPHHAAAFGTLAASVAAPGSGRVLYAVF SRDSRSSGGMGAGLCLPJDEWHAKMEANN ACYTGTREARDIFYKPFHGDIQCGGHAPOSSK SFCGSELPYPLGSRGJRGTAVLQRGGIN LTAYTVAAENNTYALGTSDGRILKVYLTP DGTSSEVDSLVENKYKRDLVLSGOLGSLY AAMTODKVFRLFVQECLSYPTCTOCRDSQDFY COWCVVGGCTEKAECFRAEASHW.WSR XSCVAYTSAQPONMSRAQGEVQLTVSPLPA LSEEDELLCFGSEPPIPARVEGAVCNSPSS IPVTPCQDHYAVTIQLLLRRGNIFLTSYQYFF YDCXQAMSLEBLLCTGSCFPIPARVEGAVCNSPSS IPVTPCQDHYAVTIQLLLRRGNIFLTSYQYFF YDCXQAMSLEBLLCTGSCVSNRWTCOWDLR YHECRAEASNNPEGDVRAHMEDSCPQCLGFSP LVPPMHETDVNRQGKNLDTVKGSSLHVGSS LKFMEVTVMGSSGTFARTFTKLSHDANETL FLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SCCAANAPDYRCAWCGGGRCVVGALCNITT SECPPYVTRQFFTGPFTGFTGFTGFTARTFTKLSHDANETL FLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC VRVTLNGYPCKVTKFGAQLQCVTGPQATRG VRVTLNGYPCKVTKFGAQLQCVTGPQATRG VRVTLNGYPCKVTKFGAQLQCVTGPQATRG VMLLEVYGGSPSVTVALCHTTTCSGED VRVTLNGYPCKVTKFGAQLQCVTGPQATRG VMLLEVYGGSPSVTVALCHTTCSGED VRVTLNGYPCKVTKFGAQLQCVTGPQATRG VMLLEVYGGSPSVTPAGRATTHOTHLDTGSGED VRVTLNGYPCKVTKFGAQLQCVTGPQATRG VMLLEVYGGSPSVTPAGTTTTGGVKCAVNGLIRAR GTDLNKAMTLQEAEAFVGARCTMKTLTET DLYCPPFVQPTTARTFTGGVKCAVNGLIRAR GTDLNKAMTLQEAEAFVGARCTMKTLTET DLYCPPFVQPTTARTFTGGVKCAVNGLIRAR GTDLNKAMTLQEAEAFVGARCTMKTLTET DLYCPPFVQPTTARTFTGGVKCAVNGLIRAR GTDLNKAMTLQEAEAFVGARCTMKTLTET DLYCPPFVQPTTARTFTGGVKCAVNGLIRAR GTDLNKAMTLQEAEAFVGARCTMKTLTET DLYCPPFVQPTTARTFTGGVKCAVNGLIRAR GTDLNKAMTLQEAEAFVGARCTMKTLTET DLYCPPFVQPTTARTFTGGVKCAVNGLIRAR GTDLNKAMTLGAACAVTLARTFTTARTFTHATTBLUSGCDD VEXTADOYCGGGRCVCAGACTYCKTRAGAC VPCATAGATA V						Į.	PNPKLEK/PVLLSEFFIKIFKGELDIVIF VENGDIK							
AKILSHPT.STHLOSSMARTGESSSECKAPLI GGRTSPYSSNGTSYVYTYSGDPPILKFKAPI EEMEEKVHOCCRIS AGPLEPRYGGMALQLWALTLLGLIGAGAS LRPRKLDFFRSKELNHLAVDEASGVVYJGA VNALYQLDAKIQLEQQVATGPVLDMKKCTP PEASQCHEAEMTDYNYQLLLVDPPRKRLYP CGQLLKGRCALRALSNIS.RLFYEDGSGEKSF VANDEGGVATGFULDNKCCTP PEASQCHEAEMTDYNYQLLLVDPPRKRLYP CGQLKGRCALRALSNIS.RLFYEDGSGEKSF VANDEGGVATGFULDNKCCTP PEASQCHEAEMTDYNYQLLLVDPPRKRLYP CGQLKGRCALRALSNIS.RLFYEDGSGEKSF VANDEGGVATGFULDNKCCTP PEASQCHEAEMTDYNYQLLLVDPPRKRLYP CGQLKGRCALRALSNIS.RLFYEDGSGEKSF VANDEGGAVTGFUNSTOFGGORVLYVGKG NGPHDMGIIVSTRLDRTDSREAFEAYTDHAT YKAGYIS.STNTQGVAAFDGGYVTYFVNQD KHPARNRTLLARMCREDPNYYSYLEMDLQC RDPDHAAAFGTCLAASVAAPGSGRYLYAVF SRDSRSSGGFGAGLCLFPLDEVHAKMEANRN ACYTGTREARDIFYEPHGOIVCGGHAPOSK SFPCGEEILPYPLGSRDGLRGTAVLQRGGLN LTAYTVAAENNHTVALGTSDGILKVYLTP DGTSSEYDBLUENRKVKRDLVLSGDLGSLY AMTQDKVFRLPVQECLSYPTCTOCKDSOPP CGWCVVGGRCTRKAACCRAEEASHWLWSRS KSCVAVTSAQRQNMSRAQGEVQLTVSPLP LSEEDELLCFGSSPPHARVEGGAVCNSPSS IPVTPPGQDHVAVTQLLLRRGNIFLTSYQYFF YDCRQAMSLEENLFCSCVSNRWTCQWDLR YHECRASPNPEGGIVRAHMEDSCPG-LGFS LIFMMETDVNPGGANLDTVKGSSLHVGSDL LLKFMEPVTMQESGTFARFIFKLSHDANETL PHLYVKSYGKNIDSKLHVLTYLOSFGRSDC SLCRAANPDYRCAWCGGGSRCVYEALCNTT SCCPPYTHIOPETGHGGGGRTHLLGSNLGVQ AGDIQRISVAGRCCSFPREYSYSTRIVCVIEA ABTIPFTGGVVFUDVTGCGGGRTHLTGSNLGVQ AGDIQRISVAGRCCSFPREYSYSTRIVCVIEA ABTIPFTGGVVFUDVTGCGGGRTHLTGNLGVG VMTLNOVPCKVTKFGQQUCYTGPQATRG QMLLEVYGGSVTPNFGFFTTREPWLRAFE PLRSFASGGRSINVTGQGFSLGRFAMVVLAEP PLRSFASGGRSINVTGQGFSLGRFAMVVLAEP PLRSFASGGRSINVTGGGFSLGRFAMVVLAEP PLRSFASGGRSINVTGGGFSLGRFAMVVLAEP CAGASETYPPTTEMFTGGVKCQVMDYGMTGKL DPEPRRPVVEQALTYSTIVQDEGGVARLINGSETVV ERMLSNWMSICLYQ'HLDSKSFLINFHIT LIENQFFSAARKYYTSUSLIPLIVTYPM VYVIAVSYYCYWKSQAGAEETEKKSQLEG LESSYDRCKKEFTDLMEMBEDQTNDVHEAG GPULDYKTYTDRYFTPEKGODECOWMTGKL DPEPRRPVVEQALTYSTIVQDEGGWALLRAR GTNLNKMSTICLYQ'HLDSKSFLINFHIT LIENQFFSAARKYYTALHGKLEYYT DMHTHLELLEGVYVACNFRLMLRRSETVV ERMLSNWMSICLYQ'HLDSKSFLINFHIT LIENQFFSAARKYYTLAHRKLEYYT DMHTHLELLEGVYVACNFRLMLRRSETVV ERMLSNWMSICLYQ'HLDSKSFLINFHIT LIENQFFSAARKYYTLAHRYNTLMHYNLYD DHTHLIELLEGVYCOPGCGGRPHALL EENRYWHYNGGGFSDAGCFTEKKINGLGG VK	İ)			,	ļ	APVSAYQYALANGDVWKVHEVPDISMAIG							
GGRTSPYSSNOTSVYTTVTSGDPPLLKFKAPI EMMEER VHGCCUS 838 2188 A 6865 6291 739		1	1		1	1	NPGVADATPPWSSYKEQSPQILLELKKQKAA							
EBMEEK JHGCCKIS AGPLEPRVOGAMALQL WALTILIGLIGAGAS LRPRKLDFFRSEKELNHLAVDEASGVYVIGA VANALYQLAKLQLEQWATOPYLDMKKCTP PEASQCHEAEMTDNYNQLLL VDPPRKRIVE CGQLLKGRCALRALSHISL RI-PEUGSGEKSF VASNBEGVATVGL SYSTGPGGDRVLFVGKG NGPHDNGIVSTKLLDRIDSREAFEAYTDHAT YKAGYLSTNTQCPV AAFEDGPYYFYLPMQDD KHPARNETLLARMCREDPNYYSYLEMDLQC RPDDIHAAAFGTLAASVAAPGSGRVLYAVF SRDSRSSSGPGAGLCLFPLDEVHAKMEANEN ACYTGTREAADIFYRFPHDQVGCGHAPOSSK SFPCGSEHLPYPLGSRDGLRGTAVLQRGGIL LTAYTVAENNHTVAFGTSDGRUKYLTP DGTSSEYDSIL VERKRYKRDLVLSGDLGSLY AMTODVERLPVOGCLSYPTCTCRDSODPY CGWCVVEGRCTRKACEPRAEBASHWLWSIS SKSCVAVTSAQPQMMSRRAQGEVQLTVSPLPA LSEEDELLCLEGESPPHPARVEGAVICNSPSS IPVTPPQQDHVAVTIQLLRRGNIELTSY QYBY YDCRQAMSLEENLPCISCVSNSWTCQWDLR YHECREASPNPEDGIVRAHMEDSCRQFLOFSS LVFNMEIDTVNFQGKNLDTVKGSFRSDL LLKFMEPVTMQESGTFAFRTPKLSHDANETL LHLYVKSYGKNDSKLHVTLUDSGRSDL SLCRAANPDVRCAWCGGGRCVYSELLCNT SCCPPPYTTRIQFETQLGGGRTILIGSNLGVQ AGDIORISVAGRNCSPOFRSYSTRIVCUE AETPTTGVSVEDVFCKLGRSPSVSTRIVCUE AETPTTGVSVEDVFCKLGRSPSVSTRIVCUE AETPTTGVSVEDVFCKLGRSPSVSTRIVCUE AETPTTGVSVEDVFCKLGRSPSVSTRIVCUE VRYTLNGVPCKYTKFGAQLCCVTPQATRG QMLLEVSVGSSPVPNFGIFFTYRENPVLRAFE PLRSFASGGRSINVTGQGDSLQRFAMVVAEP LOSWGPREABSLQPMTVGGTDVFTNDTK VYSLSPAVPEEPEAYNLTVLIEMGGHBALTR EAGAFEYVPDTTEPHFTGVFKKQVNKLIRAR GTNLNKAMTLQEABAFVGARCTMKLTINLPETIVKF GSEEWLGREFSPVFTLPSGGGKKCVNKLIRAR GTNLNKAMTLQEABAFVGARCTMKLTINLPETIVKF GSEEWLGREFSTRYSLPSCRGDKOVMTIGNLFUT DLYCEPEVQPTPREPRISTGVKKQVNKLIRAR GTNLNKAMTLQEABAFVGARCTMKLTLITLPTIVKF GSEEWLGREFSTRYSLPSGGDKOVMTIGNLFUT DLYCEPEVQPTFLPSKGDKOVMTIGNLFUT DLYCEPEVQPTFLPSKGDKOVMTIGNLFUT DLYCEPEVQPTFLPSKGDKOVMTIGNLFUT DLYCEPEVQPTFLPSKGDKOVMTIGNLFUT DLYCEPEVQPTFLPSKGDKOVMTIGNLFUT LENGPESRARKVTTASLITVALHGKLEFYT DLYCEPEVQPTFLPSKGDKOVMTIGNLFUT DLYCTPTTRYTTPLTSVTVLIEMGGHBALTD VYLEPALTYSVTVLORGAGPLYNLIKER FUT DLYCTPTTRYTTPLTSVTVLORGAGPLYNLIKER FUT DLYCTPTTVTVTVTVLICHGGLEDV VEXPLITYSVTVQDGGVADPVKVLINCDISG VKERIMOVYNGQCGCSVARPOSVLEWRPG STAQLLSDLUTSGRGEWKNTILMHTNAVL DCAATLLSKVGVGQCCSVARPOSVLEWRPG STAQLLSDLUTSGRGEWKNTILMHTNAVL DCAATLLSKVGVGQCCSVARPOSVLEWRPG STAGLLSDLUTSGRGEWKNTILMHTNAVL DCAATLLSKVGTLOOPTONFFG		ì]				AKLLSHPFLSTHLGSSMARTGESSSEGKAPLI							
838 2188 A 6865 6291 739 AGPLEPRVQGAMALQLWALTLIGLIGAGAS I RPRKLDFRSREELRIMLADEASGVYLIGA VNALYQLDAKLQLEQQVATGPVLDNKKCIP PIEASQCHEABMTDNIVQLLLVDPRKIT.PE CQLLKGICARALSNISLRLFYEDGSGEKSF VASNDEGVATVGLVSSTGRGGDRVLFVGKG NGPHDNGINISTRLDRTDRSREAFRAYTDHAT YKAGYLSTNIQQFVAAFEDGPVYFVFNOQD KIPARANTILARMCEDPNYSYLEMDLQC RDPDIHAAAFGTCLAASVAAPGSGRVLYAV SRDSRSSGGPGGLCLFPLDEVHAKMEANEN ACYTGTREABDIFVKPFHGDIQCGGHAPOSSK SFPCGSEHLPPLIGSBGDLGKGTAVLQRGGLN LTAVTVAAENNHTVAELGTBOCRLKVYLTP DOTISSEVDISL VEINKRVKRDLVLSGDLGSLY AMTODKVRLPVGECLSYPTCTQCRDSQDPY CGWCVVEGRCTRKAECPRAEEASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSRLPA LSEEDELLCLFGESPPHPARVEGEAVICNSPS PLYFPFQOHVAVTIQLLLRGRIFLTSYQYFF YDCRQAMSLEENLPGISCVSNRWTCQWDL YHECZEASPMPEDGVRAIMEDSCPQFLGPS LVFWMHEIDVRGGKNLDTVKGSSLHVGSSLHVGLWARD AGNOGRSGVARVACH ACHIEV SCHEMPTHOGESGTFAETPTLISHDANETL PLHLYVKSYGKNIDSKLHVTLVDCSFGRSDC SLCRAANFDYRCAWCGGGRCVYEALCNIT SCCPPFVTRQFGFGLGGGGTTLIGSNLGVQ AGDIQRISVAGRNCSPOPERTSYSTRIVCVTEA AETPTTTELTCHTLGSNLGVG AGDIQRISVAGRNCSPOPERTSYSTRIVCVTEA AETPTTTELTCHTLGSNLGVG KPLSVEPQQOPQAGGTTLTHGTHLDTGSGCD VRVTLNGVPCKVKFGAQLQCVGTQQATRG QMLLEVSYGGSPVPNGIGFTTVRENPVLRAF PLRSSAGGRSINVTGGOPSLGVGRAMVIAEF PLRSSAGGRSINVTGGOPSLGVGRAMVIAEF PLRSSAGGRSINVTGGOPSLGVGRAMVIAEF PLRSSAGGRSINVTGGOPSLGVGRAMVIAEF PLRSSAGGRSTNVTGGTSVTPFNVTLARG GSEWVLGRVETOTRYSDYPLSLIPLIVVFN VVVANSYYCVWRKSQQAFETEXFEKSQLEG LESSVRDKKKEFTDLMEMEDQTINDVHEAG IPVLDXYTTDRYFLENSDGJGKDVMITIGAL INDXXTYTTGTVFFLENSDGLAGHUNDSTATURGTDLGDD VEYAPLTVSVTVQDGAVARGKETVNLINLFINVERG FRAMSNINGLAGAPVYPDTTERFTGGVKKQVMKLIRAR GTINLNKAMTLQEABAPVYEQALVGPSVLLMRRSETVV GREWNLGRVFTDTSVTLRAGGLECTVV DDMHTLFLELLGQVVVAXVPKLMRRSETVV GREMLSNWMSICLYQYTLGSAGPLYNLKEH LNENGPEFSARAVYTASJLTVALIGKLEFYT DDMHTLFLELLGQVVVAXVPKLMRRSETVV GREMLSNWMSICLYQYTLGSAGPLYNLTWINDTIGLGDD VEYAPLTVSVTVQDGAVARAVTNILDTIGLGDD VEYAPLTVSVTVQDGAVARAVTNILDTIGLGDD VEYAPLTVSVTVQDGAVARAVTNILDTIGLGDD VEYAPLTVSVTVQDGAVARAVTNILDTIGLGDD VEYAPLTVSVTVQDGAVARAVTNILDTIGLGDD VEYAPLTVSVTVQDGAVARAVTNILDTIGLGDD VEXAPLTYSVTVQDGAVARAVTNILDTIGLGDD VEXAPLTYSVTVVQDGAVARAVTNILDTIGLUGD VALETENT THE DATTLINGTLOGOPTO	Ì	1	ì		İ		GGRTSPYSSNGTSVYYTVTSGDPPLLKFKAPI							
LRPRKLDFFRSKELNHLAVDEASGVYVLGA VNALYQLDAKLQLEQOVATOPYLDNKCTP PIEASQCHEAEMTDNVNQLLLVPPPRKRIVE CQQLLKGICALRALSINSILLFPEDGSGEKSF VASNDEGVATVGLVSSTGPGDRVLFVGKG NGPHDNGIVSTRLLDRTDRSREAFRAYTDHAT YKAGYLSTNTQPVAAFEDGPYVFFVNQOD KHPARNETLLARMCREDPNYYSYLEMDLOC RDPDHHAAAFGTLAASVAAPGSGRVLYAVF SRDSRSSGGPGAGLCLFPLDEVHAKMEANEN ACYTGTREARDFYKFPHDDQCGGHAPOSK SFPCGSEHLPYPLGSRDGLRGTAVLQRGGLN LTAVTVAAENHTVALGTSTDGRILKVYLTP DGTSSEYDSILVENKRVKRDLVLSGDLGSLY AMTQDAVSTRLYDGCLSYPTCTGCBSQDPY CGWCVVGGCCTRKACCPRAEBASHWLWSIS KSCVAVTSAQPOMSRRAQGEVQLTVSPLA LISEDELLCLFGESPPHRAVEGAVICNSPS IPVTPPGQDHVAVTIQLLRRGNIFLTSYQYFP YDCRQAMSLEENLFCISCVSNRWTCOWDLR YHECREASPNPEDGTVAAHMEDSCTQFLOPS LIKPMEPTVINQESGTFAFTFRLSHDANETL LHLYVKSYGKNDSKLHVTLVDSFGRSDC SLCRAANPDVRCAWCGGGRCVYBALCNTT SECPPPYTTRQFGPLGGGGTTLINGSLOVQ AGDIQRISVAGRNCSFQFRTSYSTRIVCVUE AETPFGGVEUDYGKLGRSFPNVQFTFQQA KPLSVEDQQPQAGGTTLTHGTHLDTGSGCA AETPFTGGVEUDYGKLGRSFPNVQFTFQQA KPLSVEDQQPQAGGTTLTHGTHLDTGSCG WLLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQA AETPFTGGVEUDYGKLGRSFPNVQFTFQAF KPLSVEDQQPQAGGTTLTHGTHLDTGSCG WLLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF GMLLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLSVSTGGSPVPNFGIFTYRENPVLRAFE PLRSTASGGRSINVTQQGFGCAWRAGARCTINKTLTTT DLYCCPPEVQPPTRSTRGGVCKACVTNALPETUKFGAF GSREWLGRAFAYASLTAVLHGRAGLEYT UNDMTTLTBLELDQFVVARNFKLLHYTT DLYCTPTVQFTLNSPCGGCSCWRRDSVVLEWRG STAGLLSSLUTTGRGGRWKNVTLMHTYNVL DGATTLLSKYGTLOGPTONFGSVL		ļ.	1				EEMEEKVHGCCRIS							
LRPRKLDFFRSKELNHLAVDEASGVYVLGA VNALYQLDAKLQLEQOVATOPYLDNKCTP PIEASQCHEAEMTDNVNQLLLVPPPRKRIVE CQQLLKGICALRALSINSILLFPEDGSGEKSF VASNDEGVATVGLVSSTGPGDRVLFVGKG NGPHDNGIVSTRLLDRTDRSREAFRAYTDHAT YKAGYLSTNTQPVAAFEDGPYVFFVNQOD KHPARNETLLARMCREDPNYYSYLEMDLOC RDPDHHAAAFGTLAASVAAPGSGRVLYAVF SRDSRSSGGPGAGLCLFPLDEVHAKMEANEN ACYTGTREARDFYKFPHDDQCGGHAPOSK SFPCGSEHLPYPLGSRDGLRGTAVLQRGGLN LTAVTVAAENHTVALGTSTDGRILKVYLTP DGTSSEYDSILVENKRVKRDLVLSGDLGSLY AMTQDAVSTRLYDGCLSYPTCTGCBSQDPY CGWCVVGGCCTRKACCPRAEBASHWLWSIS KSCVAVTSAQPOMSRRAQGEVQLTVSPLA LISEDELLCLFGESPPHRAVEGAVICNSPS IPVTPPGQDHVAVTIQLLRRGNIFLTSYQYFP YDCRQAMSLEENLFCISCVSNRWTCOWDLR YHECREASPNPEDGTVAAHMEDSCTQFLOPS LIKPMEPTVINQESGTFAFTFRLSHDANETL LHLYVKSYGKNDSKLHVTLVDSFGRSDC SLCRAANPDVRCAWCGGGRCVYBALCNTT SECPPPYTTRQFGPLGGGGTTLINGSLOVQ AGDIQRISVAGRNCSFQFRTSYSTRIVCVUE AETPFGGVEUDYGKLGRSFPNVQFTFQQA KPLSVEDQQPQAGGTTLTHGTHLDTGSGCA AETPFTGGVEUDYGKLGRSFPNVQFTFQQA KPLSVEDQQPQAGGTTLTHGTHLDTGSCG WLLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQA AETPFTGGVEUDYGKLGRSFPNVQFTFQAF KPLSVEDQQPQAGGTTLTHGTHLDTGSCG WLLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF GMLLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLSVSTGGSPVPNFGIFTTYRENPVLRAFE PLRSTASGGRSINVTQQGFSLQRFAWVGTFQAF CMLSVSTGGSPVPNFGIFTYRENPVLRAFE PLRSTASGGRSINVTQQGFGCAWRAGARCTINKTLTTT DLYCCPPEVQPPTRSTRGGVCKACVTNALPETUKFGAF GSREWLGRAFAYASLTAVLHGRAGLEYT UNDMTTLTBLELDQFVVARNFKLLHYTT DLYCTPTVQFTLNSPCGGCSCWRRDSVVLEWRG STAGLLSSLUTTGRGGRWKNVTLMHTYNVL DGATTLLSKYGTLOGPTONFGSVL		- 2100	+	4065	6201	730	AGPLEPRVOGAMALQLWALTLLGLLGAGAS							
WALYQLDAKLQLEQQYATGPVLDNKKCTP PERASCCHREARTDNYNQLLLVDPPRKRI.VE CQCILKGRCALRALSNISRLFYEDGSGEKSF VASNDEGVATVGLVSTOFGGERVLFVGKG NOPHDNGIVSTRLLDRTDSREAFEAYTDHAT YKAGYLSTNTQQFVAAFEDGPYVFFVFNQDQ KHPARNRTLLARMCREDPYYSYLEMDLQC RDPDHAAAFGTCLAASVAAPGSGRVLYAVF SRDSRSSGGFGAGLCLFPLDEVHAKMEANRN ACYTGTREARDIFYKPHGDIQCGGHAROSSK SPPCGSEHLPYPLGSRGULGTAVLQRGGLN LTAVTVAAENNHTVAFLGTSDGTLKVYLTP DGTSSEYDSILVENREVKRADLVSGDLOGSLY AMTQDKVFRLPVQSCLSYPTCTQCRDSQDPY CGWCVVCGRCTRKAECPRAEASHWLWSRS KSCVAVTSAQFQNMSRRAQGEVQLTVSPLPA LSEEDELLCLFGESPPIHARVEGEAVICNSPSS IPVTPFQQDIPVAVTIQLLERRGNIELTSYQYPF YDCRQAMSLEENLPCISCVSNRWTCQWDLR YHECRASPNPEDGIVRAHMEDSCPQFLOPSP LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD LIKFMSPTVMDGSGTFAFFTKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANDYRCAWCGGGSRCVBALCNTT SECPPVTTRQFFTGPLGGGIRTTLGSNLGVQ AGDIGRISVAGRRCSFQFBYSVSTRUVIEA AETPFTGGVEVDVFGKLGSKSPNVQFTFQQP KPLSVEPQQGPQAGGTTLTHIGTHLDTGSQED VRVTLRGVPCVWFKGFGAQCVGGGRAV QMLEVSYGGSPVNPGIFFTYRENPUXAFE PLRSSASGGSSINVTQGGFSLQCVFGQATRG QMLLEVSYGGSPVNPGIFFTYRENPUXAFE PLRSSASGGSSINVTQGGFSLQCVFGQATRG QMLEVSYGGSPVNPGIFFTYRENPUXAFE PLRSSASGGSSINVTQGGFSLQCRAMVVIAEP LQSWQPRRAESLQPMTVVGTDVYFHNDTK VVFLSAVPEEPAAYLTLTLEMDGHRALKT EAGAFYVPDPIFENFTGGVKXQVNKLIRAR GTNLNKAMTLQBAEAFVGAECTMKILTET DLYCCPPEVQFPKRRQKKADTHINLFEIVKF GSREWVLGRVEYDTRVSOVPLSLIPLLVIVPM VVVIAVSYYCYVWKSQQAEACPSKIKSQLEG IPVLDKTYTTDNVFFLSKOGDKDVMTGKL DDEPRRPVVSQALYQFSKLLNSKSFLINFHH LLENGPESARAVYYASLLVILANGRLEXTV BRMLSNWMSICLYQYLKBGRAUKRASETVV BRMLSNWMSICLYQYLKBGRAUKRARSTVV BRMLSNWMSICLYQYLKBGRAUKRASETVV BRMLSNWMSICLYQYLKBGRAUKRANERSTVV BRMLSNWMSICLYQYLKBGRAUKRVNITHMYNR DGATLILSKVQSQQFESQQDLPGERHALL EEGRRAVHLVRYTDEVDEGKSKRGSVCKEE PVLATENTYROPGGROWERVSVLEWRPG STAQILSDLDLTISQCGROWKRVNITHMYNR DGATLILSKVGNGQPCSCWPRSDSVLLEWRPG STAQILSDLDLTISQCGROWKRVNITHMYNR DGATLILSKVGNGQPCSCWPRSDSVLLEWRPG STAQILSDLDLTISQCGCRWKRVNITHMYNR DGATLILSKVGNGQPCSCWPRSDSVLLEWRPG STAQILSDLDLTISQCGCRWKRVNITHMYNR DGATLILSKVGNGQPCSCWPRSDSVLLEWRPG STAQILSDLDLTISQCGCRWKRVNITHMYNR DGATLILSKVGNGQPCSCWPRSDSVLLEWRPG	838	2188	I A	6005	0231	1 127	LRPRKLDFFRSEKELNHLAVDEASGVVYLGA							
PIEASCCHEAEMTDNYNQLLLVPPPRKI.VE COQLIKGICALRALSINISHLFYEDGSGEKSF VASNDEGVATYGLVSSTOPGGDRYLFYGKG NGPHDNGIVSTRLLDRTDSREAFEAYTDHAT YKAGYLSTNTQGVAAFEDGPYVFFVPNQDD KHPARNRTLLARMCREDPYYSYLEMDLQC RDPDDHAAAFGTCLASVAAPGSGRVLYAVF SRDRSSGGFGAGLCLFPLDEVHAKMEANRN ACYTGTREARDIFYKPFHGDIQCGGHAPGSKK SFPCGSEH,PYPLGSRDGLRGTAVLQRGLN LTAVTVAAENHTVAFLGTSORLKVYLTP DGTSSEYDSILVEINKRVKRDLJVLSGDLGSLY AMTODKVFRLPVQSCLSYTCTGCDSODPY CGWCVVGRCTKKAECRAEEASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSHPA LESEDELLCIFGESPPIPARVEGGAVICNSPSS IPVTPGQDHVAVTIQLLLRRGNIFLTSVQYFF YDCRQAMSLEENLFCISCVSNFWCOWDLR YHECREASPNPEDGIVRAHMEDSCPGFLGPSP LVIPMNHETDVNFGGKNLDTVSCSLHVGSD LKFMEPYTMQESGTFAFRTPKLSHDANETL PLHLYVKSYGKNDSKLHVTLYDCSFGGASLOVQ AGDIGNSVAGRNCSFGPERSYSSTRIVCVEA AETPFTGGVEUDVFGKLGRSPPNVQFTQOP KPLSVEPQOPQAGGTTLTHGTHLOTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLEVSYGGSPVPNPGFFTYRDVLRAFE PLRSFASGGRSINVTQGFFSLIQRRAMVVAEP LQSWQPPREASSLQPMTVVGTDYFHNDTK VVFLSRAVEPEPAYNITULEDDGHRALLRT EAGAFEVVPDTFENTIGGVKKQVNKLLRAR GTNLNKAMTLGAEAFVGGKRCTHNTLTT LQSWCPPREASSLQPMTVVGTDYFHNDTK VVFLSRAVEPEGAYNCTOVERPOYTER GSREWVLGRVEYDTRSDVPLSLIPLFLVVFM WVYLSSVYCGYPPCARGCRCTMTLGSNLLRAR GTNLNKAMTLGAEAFVGGKRCTHTHNTLK LGSGEPVVPDTFENTIGGVKKQVNKLLRAR GTNLNKAMTLGAEAFVGGKRCTHTHNTLE LGSWCPPEVQPPFKRRGKBDTTHNLEFTVL GSREWVLGRVEYDTFNSDVPLSLIPLFLVKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLVKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLVKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLVKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLVKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLVKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLVKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLVKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLVKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLYKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLYKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLYKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLYKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLYKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLYKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLYKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLYKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLYKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLYKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLYKFM GSREWVLGRVEYDTRVSDVPLSLIPLFLYKFM GSREWVLGRVEGTDRDSVVLEWRFG TAGLISDLDLTSQCGGRWKRVNTIMHTNYR DGATLLISKVGGOFGCSWCMPROSVVLEWRFG STAQILSDLDLTSQCG	1	1	ļ		ł	l	VNAL YOU DAKLOLEOOVATGPVLDNKKCTP							
CGQLLKGICALRALSNISLRLFYEDGSGERSF VASNDEGWATVGLVSTOFGGGRVLFVGKG NGPHDNGIIVSTRLLDREAFEAYTDHAT YYAGYLSTNTQGYVAAFEDGYVYFYFVNQDD KHPARNRTLLARMCREDPYYYSYLEMDLQC RDPDHAAAFGTCLAASVAAPGSGRVLYAVF SRDSRSSGGFGAGLLFFIDEVHAKMEANRN ACYTGTRARDIFYXFYHGDIQCGGHAFOSK SFPCGSEHLPYPLGSRDGLRGTAVLQRGGLN LTAVTVAAENNHTVAFLGTSDCRILKVYLTP DGTSSEYDSILVEINKRVKRDLVLSGDLGSLY AMMODKVFRLFVQBCLSYPTCTQCRDSQDPY CGWCVVGRCTTKAAECPRAEEASHW.WSRS KSCVAYTSAQPOMNSRRAQGEVQLTVSPLPA LSEEDELLCLFGGSPPIPARVEGGAVLCNSPSS IPVTPPQQDIVAVTQLLLRRGHIELTSYQYFP YDCRQAMSLEENLPCISCVSNRWTCQWDLR YHECRASPNTEDGIVRAHMEDSCPPLGPSP LLVFMNHEIDTWNPQGKNLDTVKGSSLHVGSD LLKFMEPVTMQGSGTAFRTFKLSHDAMETL PHLLYVKSYGNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGGSRCVVEALCNTI SECPPPYTRIQPFTCPTGGGGRTLGNSLGVQ AGDIQRISVAGRNCSFQPERYSVSTRIVCVUEA AETPTTGGVEVDVFGKLGRSPPNVQFTFQQP KPLSVEPQQGPQAGGTTLTHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGFGATRG QMLLEVSYGGSVPNPGIFTYRENPULAAFE PLSSFASGGSINVTQGGFSLQRFAMVVIAEP LQSWQPPREAESLQPMTVVGFTYRENPULAAFE PLSSFASGGSSINVTQGGFSLQRFAMVVIAEP LQSWQPPREAESLQPMTVVGFTYRENPULAAFE PLSSFASGGRSINVTQGGFSLQRFAMVVIAEP LQSWQPPREAESLQPMTVVGFTYRENPULAAFE GGREWULGRVSYDTYRSDVYLSLILLPLIVVPM VVYLAVSVYCYWEKSQQAEREYEKKLRAR GTNLNKAMTLGAEAFVGAGRECTMKTLTET DLYCEPPEVQPTPFRRPRVLSALF GSREWULGRVSYDTYRSDVPLSLILLPLIVVPM VVYLAVSVYCYWEKSQQAEREYEKKSQLEG LEESVRDRCKKEFTDLMEMGEDOTNDVIEGG BPVLDYKYYTDRVFFLPSKDGDKDVMITGKL DIPERRRVVEQALYQFSNLLNSKSLINHTH LENGPESSARAVYTASLYLAKGKLEYY DDMTLFLELLEQYVVAKNFKLMLRRSSTVV REMS.SNWMSICLYQYLKDSAGGPLYKLFKAI KHQVEKGPVDAVQKKAYYTLNDTGLGDD VEYAPLTVSIVULDGEVGROPSVVLEWPG STAQILSDLDLTSQRGRWKRVNTILMTHYNVR DGATLLISKVGVGSQQFEDSGARSTVKKENT KHQVEKGPVDAVQKKAYYTLNDTGLGDD VEYAPLTVSIVULDGEVGROPSVVLEWPG STAQILSDLDLTSQRGRWKRVNTILMTHYNVR DGATLLISKVGVGSQQFEDSGOQDLPGERHALL EEGRRVWHVLVRPTGEVGEVSKCRVSKEKE PVX ATTEVT TELL SLYKGTLOGFDJONFGSVKL	Į.	l.	Į .	1		l .	PIEASOCHEAEMTDNVNOLLLVDPPRKRLVE							
VASNDEGVATVGLVSSTGPGGDRVLFVGKG NGPHDNGIVSTGLDRIDSRAFEAYATDHAT YKAGYLSTNTQQFVAAFEDGPYVFFVQDD KHPARNRTLLARMCREDPHYYSYLEMDLOC RDPDIHAAAFGTCLAASVAAPGSGRVLYAVF SRDSRSSGGGGAGLCLFPLDEVHAKMEANRN ACYTGTREARDFYKPFHGDIQCGGHAPGSK SFPGGSEHLPYPLGSRDGLRGTAVLQRGGLN LTAVTVAAENHTVAFLGTSDGRLKVYLTP DGTSSEYDSILVENNRKVRDLVLSGDLGSLY AMTODLVFRLPYQECLSYPTGCRDSODPY CGWCVVEGRCTRKAECPRAEEASHWLWSRS KSCVAVTSAQPQNNSRRAQGEVQLTVSPLPA LSEEDELLCLFGESPPHARVEGEAVICNSPSS IPVTPGQDHVAVTIQLLLRRGNIFLTSYQYPF YDCRQAMSLEERLPCISCVSNRWTCQWDLR YHECREASPNPEDGIVRAHMEDSCPQFLOPSP LVIPNNHETDWFOGNLDLTVGSSLHVGSD LLKFMEPVTMQESGTFAFRTFKLSHDANETL PLHLYVKSYGKNDSKLHVILYDCSGGRSDC SLCRAANPDYRCAWCGGGSRCVVEALCNTT SECPPFVTRIQPETEUFLGGBRTHLGSNLOVQ AGDIQRISVAGRRCSSPQERSYSVSTRIVCVIEA AETPFTIGGVEUDVFGKLGRSPPNVQFTFQOP KPLSVEPQQGPQAGGTLTLIHGTHLDTGQQE VRVTLNGVPCKVTKFGAQLQCVTGPAJRG QMLLEVSYGGSPVPPPGFTFYRENPULRAFE PLRSFASGGRSINVTQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHINTK VVFLSAVYEEPPAAYLITLUEMDGHRALLRT EAGAFEYVPDTFENFTIGGVKQVNKLILRR GTNLNKAMTLGAEAFVGGRCTVHLENDGHRALLRT EAGAFEYVPDTFFFFTGGVKQVNKLILRR GTNLNKAMTLGAEAFVGGRCTHLENDGHRALLRT EAGAFEYVPDTFFFTFTGGVKQVNKLILRR GTNLNKAMTLGAEAFVGGRCTHLENDGHRALLRT EAGAFEYVPDTFFFTFTGGVKQVNKLILRR GTNLNKAMTLGAEAFVGGRCTHLENDGHRALLRT LENGGPESVARRAVYTASLTLANDHAAC BPLDYKTYTTNOFFLFFKSGGRDDVMTGKL DIPERRPVVEQALYQFSNLLNSKSFLINFHT LENGGPESVARRAVYTASLTIVALHGKLEYV DEMTLFLELLEQYVVAKNFKLMLRRSETVV BRMLSNWMSICLYQYKLDSAGEPLYKLFNHT LUENGPETSARAVYTASLTIVALHGKLEYTV DEMTLFLELLEQYVVAKNFKLMLRRSETVV BRMLSNWMSICLYQYKLDSAGEPLYKLFSTVV BRMLSNWMSICLYQYKLDSAGEPLYKLFSTVV BRMLSNWMSICLYQYKLDSAGEPLYKLFSTVV BRMLSNWMSICLYQYKLDSAGEPLYKLFSTVV BRMLSNWMSICLYQYKLDSAGEPLYKLFSTVV BRMLSNWMSICLYQYKLDSAGEPLYKLFSTVV BRMLSNWMSICLYQYKLDGDTSQ VEYAPLTVSTVTQDEGVDAFDSVLEWRPG STAQILSDLDTTSQEGRWKRVNTILMYNVR DGATLLISKYGVGQOPESGVCKEE PVATEVTHUSTLYGGGGGROUPCGERHALL BEERRVWHLVRPTDEVDEGKSKGSVCKEE PVATEVTHUSTLYGGGGROUPCGERHALL BEERRVWHLVRPTDEVDEGKSKGSVKEEE PVATEVTHUSTLYGGGGROUPCGERHALL BEERRVWHLVRPTDEVDEGKSKGSVKEEE			1	1		1	CCOLL PCICAL PALSNISL RI EVEDGSGEKSE							
NGHDNGIIVSTRLLDRIDSREAFERATTOHAT YKAGYLSTNOQFWAREDGPYVFFVFNQD KHPARNRTLLARMCREDPNYYSYLEMDLQC RPDDHAAAFGTCLAASVAAPGSGRVLYAVF SRDSRSSGGFGAGLCLPFLDEHAAKMEANRN ACYTGTREARDIFYKPHGDIQCGGHAGSSK SPFCGSEILPYPLGSRDGLAGTAVLQRGGLN LTAVTVAAENNHTVAFLGTSDGRILKYVLTP DGTSSEYDSILVENKRYKRKDLVLSGDLGSLY AMTQDKVFRLPVQECLSYPTCTQCRSQDPY CGWCVVEGRCTRKAECPRAEASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLCLFGESPPHPARVEGEAVLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLLFGESPPHPARVEGEAVLWSPS IPVTPRQDHAVTIQLLLRRGNIFLTSYQYFF YDCRQAMSLEENLPCISCVSNRWTCOWDLR YHECREASPNPEDGIVRAHMEDSCPQFLOPSP LVFMMHETDVNFQGKNLDTVKGSSLHVGSD LLKFMEPVTMQESGTFAFTFLSLEIDANETL PLHLYVKSYGKNIDSKLHVTLYDCSSGRSDC SLCRAANPDYRCAWCGGGSRCVYEALCNIT SECPPVTITIQPETGPLGGGRITILGSNLGVQ AGDIQRISVAGRNCSPQFREYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP KPLSVEPQQGPQAGGTTLTHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGQATRG QMLLEVSYGGSPVPNSGFTYKENPVLRAFE PLRSFASGGRSINVTGQGSLQRFAMVVLAFE PLRSFASGGRSINVTGQGSLQRFAMVVLAFE PLRSFASGGRSINVTGQGSLQRFAMVVLAFE PLRSFASGGRSINVTGQGSLQRFAMVVLAFE PLSSFASGGRSINVTGQGSLQRFAMVVLAFE PLSSFASGGRSINVTGQGSLQRFAMVVLAFE PLSSFASGGRSINVTGQGSLQRFAMVVLAFE PLSSFASGGRSINVTGQGSLQRFAMVVLAFE PLSSFASGGRSINVTGQGSLQRFAMVVLAFE PLSSFASGGRSINVTGQGSLQRFAMVVLAFE PLSSFASGGRSINVTGQGSLQRFAMVLAFE PLSSFASGGRSINVTGQGSLQRFAMVILAFT EAGAFEVYPPFTFENFTGGVKKQVNKLLRAF GTNLNKAMTLGBAEAFYGARCTMKTLTET DLYCEPPEVQPPFRRRQKGDTINHPEFIVKT GSREWVLGNEVDTURSDVPLSLILPLVTVPM VVVIAVSVYCYWRKSQQAEREYSKIKSQLEG LEESVRDRCKKEFTDLMEDEFIVKT GSREWVLGNEVDTURSDVPLSLILPLVTVPM VVVIAVSVYCYWRKSQQAEREYSKIKSQLEG LEESVRDRCKKEFTDLMERSSTVV BEMLISNWMSICLYQVLKDSAGGFLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLGDD VEYARLTVSVIVQDEGVDAPVKVLNCDTISQ VEKEINDQVYRGGPCSGWRCDNSVLEWRPG STAQLSDLDLTSGREGRWKRVNTLMHTYNK DGATLLISKGVGSQGPEGQDLVCERHALL EEENRWHHJVRTDEVDEGKSKRGVKEKER PXX ATTERTJ TILL SVKGTQLOOFVDNFFGSVL	1	1	1			1	COOLERGICALISATION TO THE TOTAL TO THE TOTAL TO THE TOTAL THE TOTAL TOTAL TO THE MTQDKVFRLPVQCCLSYPTCTQCRDSQDPY CGWCVVEGRCTREACCPRAEASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLCJEGSEPPHARVEGEAVICNSPSS IPVTPPQQDHVAVTIQLLLRRGNIFLTSYQYFP YDCRQAMSLEENLPCISCYSMRWTCQWDLR YHECREASSNPEDGIVRAHMEDSCPQFLGPSP LVIPMMHETDVNFQGKNLDTVKGSSIHVGSD LLKRMBEVTMQESGTFARTFKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAAMPDYRCAWCGGRSCVVFALCNTT SECPPPVITRIQPETGPLGGGRITTLGSNLGVQ AGDIQRISVAGRNCSFQFERYSVSTRIVCVIEA AETFFTGGEVDVPFCKLGRSPRNVQFTFQQP KPLSVEPQQOPQAGGTTLTHGTHLDTGSQED VRVTLNGBYCKVTKFGAVGCVTGPQATRG QMLLEVSYGGSPVPNFGIFTYRENPVLRAFE PLRSFASGGRSINVTQGFSLIQRFAMVVIAEP PLRSFASGGRSINVTQGFSLIQRFAMVVIAEP LSWQPPREAESLQPMTVGTTYPRINDTK VVFLSPAVPEEFFEAYNLTVLIEMDGHRALLRT EAGAFEY YPDFTFENTIGGVKKQVMKLIRAR GTILMKAMTLQEAEAFYGABRCTMKILTET DLYCEPPEVQPPPKRRQKRDTTINLPEFIVKF GSREWVLGRVFYDTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYSKIKSQLEG LEESWRDRCKKEFTDLMEIMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDOMTIGKL DIPPERRPVVEQALVQFSILLNSKSFLINFHT LENQPEFSARAKVYTASLLTVALHGKLEYYT DMHTLFIELLECYVVAKNPKLMIRRSETVV ERMLSNWMSICLYQYLURSAGEPLYSLIKKAI KHQVEKGPVDAVQKKAKYTLNDTSLQLGD VEYAPLTVSVIVQDEGGVARPKVNICDTISQ VEKEIDQVYRGQPCSCWPRPDSVVLEWPG STAQLSDLDLTSQREGRWKRWTILMTYNY DGATLLISKGVGSQCPCSCWPRDSVVLEWPG STAQLSDLDLTSQREGRWKRWTILMTYNY DGATLLISKGVGSQCPCSCWPRDSVVLEWPG STAQLSDLDLTSQREGRWKRWTILMTYNY DGATLLISKGVGSQCPCSCWPRDSVVLEWPG STAQLSDLDLTSQREGRWKRWTILMTYNY DGATLLISKGVGSQCPCSCWPRDSVVLEWPG STAQLSDLDLTSQREGRWKRWTILMTYNY DGATLLISKGVGSQCPCSCWPRDSVVLEWPG STAQLSDLDLTSQREGRWKRWTILMTYNY DGATLISKGVGSQCPCSCWPRPGSVLEWFG STAQLSDLDLTSQREGRWKRWTILMTYNY DGATLLISKGVGSQCPGDSQQDLPGERHALL EEENRWHUVRTDEVDEGKSKRGSVEKEK	1		1		1	1	VASINDEGVAT VOLVSSTOTOODAVEL VORO
KHPARNRTLLARMCREDPNYYSYLYADF RDDHAAAFGTCLASVAPGSGRVLYAVF SRDSRSGGGAGLCLFPLDEVHAKMEANRN ACYTGTREARDIFYKPFHGDIQCGGHAPGSSK SFPCGSEILPYPLGSRDGLRGTAYLQRGGLN LTAYTVAAENNHTVAFLGTSDGRLKYYLTP DGTSSEYDSILVENKRYKRDLVLSGDLGSLY AMTQDKVFRLPVQECLSYPTCTQCRDSQDFY CGWCVVEGGCTRFAECFRAEASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSFLPA LSEEDELLCLFGESPPHPARVEGEAVICNSPSS IPVTPRQQDHVAVTIQLLRRGNIFLTSYQYFF YDCRQAMSLEENLPCISCVSNRWTCOWDLR YHECRASFNPEDGIVRAHMEDSCPQFLGPSP LVIPAMHETDVNFQGKNLDTVKGSSLHVGSD LLKFMEPVTMQESGTFAFTFRTSJHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGQSRCVYEALCNIT SECPPPVTRIQPETGPLGGGRITITLGSNLGVQ AGDIQRISVAGRNCSFQFERYSVSTRIVCVIEA ABTFFTGGVEVDVFGKLGRSPFNVQFTFQOP KPLSVEPQQGPQAGGTTLTHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNGFTYRENPVLRAFE PLRSFASGGRSINVTQGGTSLQRFAMVIAEF PLRSFASGGRSINVTQGGSLQVFAMVIAEF PLRSFASGGRSINVTQGGTSLQRFAMVIAEF PLRSFASGGRSINVTQGGSLQRFAMVIAEF PLRSFASGGRSINVTQGGTSLQRFAMVIAEF PLRSFASGGRSINVTQGGSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAEF PLRSFASGGRSINVTQGTSLQRFAMVIAER PLRSFASGGRSINVTQGTSLQRFAMVIAER PLRSFASGGRSINVTQGTSLQRFAMVIAER PLRSFASGGRSINVTQGTSLQRFAMVIAER PLRSFASGGRSINVTQGTSLQRFAMVIAER PLRSFASGGRSINVTQGTSLQRFAMVIAER PLRSFASGGRSINVTQGTSLQRFAMVIAER PLRSFASGGRSINVTQGTSLQRFAMVIAER PLRSFASGGRSINVTQGTSLQRFAMVIAER PLRSFASGGRSINGTMTMTTMTYMV VVIAAVSVYCYWRKSQQAERCYKKSQLER PLRSFARAKVYTASLLTVALHGKLEYT DMHTLFILLELGQVVAKAPKLMRRSSTVV PRAMLSMWSLCLYQVLKDSAGGPLYSLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLGDD VEYANTEMTUT BLVSKVENCEN STAGLSDLLTSGREGRWKRVTILMTYNWR DGATLLSKGVSQQGPESQQDLCGERHALL EEENRYWHLVRTTDEVDEGKSKGKEKE PLTATETTUT BLVSKCTALOOFVDDFSOSVLE	1	}	1	1	I	{	NGPHDNGHV31KLLDKID3KEAFEATIDHAT							
RPDHHAAAFGTCLAASVAAPGSGRVLYAVF SRDSRSSGFOAGLCLEPLDEVHAKMEANRN ACYTGTREARDIFYKPFHGDIOCGGHAPGSK SFPCGSEHLPYPLGSRDGLRGTAVLQRGLN LTAYTVAABANHTVAFLGTSDGRILKVYLTP DGTSSEYDSILVEINKRVKRDLVLSGLGJGSLY AMTODKVFRLPVQECLSYPTCTQCRDSQDPY CGWCVVEGRCTRKAECPRAERASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLCLEGSEPPHARVEGEAVICNSPSS IPVTPPGQDHVAVTIQLLLRRGNIFLTSYQYFF YDCRQAMSLEENLPCISCYNRWTCQWDLR YHECREASPNPEDGIVRAHMEDSCPQFLGPSP LVIPMHETDVNFQGKNLDTVKGSSLHVGSD LLKFMBEVTMQESGTFARTFRLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGGGSCVVEALCNIT SECPPYTRIOPETGPLGGGRITTLGSNLGVQ AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA AETPFTGGVEVDVFCKLGRSPPNVQFTFQQP KPLSVEPQQOPQAGGTTLTHIGHTLDTGSQED VRVTLNGGYPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNFGFTYRENPVLRAFE PLRSFASGGRSINVTQQGFSLIQRFAMVVAFE LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGRALLTT DLYCEPPEVQPPKRRGKQTTHINLFPLVKF GSREWVLGRVEYDTTRVLIEMTGKRALLTT DLYCEPPEVQPPKRRGKRDTTINLFPLVKF GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAERTYSKKSQLEG LEESWRDRCKKFTDLMBEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMTGKL DIPPERRPVVEQALVQPSILLIPLVIVPM VVVIAVSVYCYWRKSQQAERTSKKSQLEG LEESWRDRCKKFTDLMBEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMTGKL DIPPERRPVVEQALVQSSLLIPLSVIVP DMHTLFIELLEQYVVAKNFKLMLRRESTVV ERMLSNWMSICLYQVLKDSAGEPLYSLKKAL KHQVEKGPVDAVQKKAKYTLNDTGLGDD VEYAPLTVSVVQDGEGVARPRENVLEVRPG STAQLSDLDLTSQREGRWKRVNTLMHYNVR DGATLLISKGVSQQCPGSCVPRPDSVVLEWRPG STAQLSDLDLTSQREGRWKRVNTLMHYNVR DGATLLISKGVSQCGPGSQQDLPGERHALL EEENRWHLVRTDEVDEGKSKGSVKEKE	1	1	1	1			YKAGYLSINIQQFVAAFEDUFIVFFVFNQQD							
SRDSRSSGOPGAGLCLFPLDEVHAKMEANRN ACYTGTBEARDIFYEPHGDIQCGGHAPGSSK SFPCGSEHLPYPLGSRDGLRGTAV1QRGGLN LTAVTVAAENNHTVAFLGTSDGRILKVYLTP DGTSSEVDSULVENKRKRKDLVLSGDLGSLY AMTODKVFRLPVQECLSYPTCTCQRDSQDPY CGWCVVEGRCTRKAECPRAEEASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLCLFGESPPIPPARVEGEAVICNSPSS IPVTPGQDHVAVTIQLLLRRGNIFLTSYQYFF YDCRQAMSLEENLPCISCVSNRWTCQWDLR YHECREASPNPEDGIVRAHMEDSCPQFLGPSP LVIPMNHETDVNPQGKNLDTVKGSSLHVGSD LLKFMEPVTMQESGTFAFRTFLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGQSRCVYEALCNTT SECPPVITRIQPETGPLGGGIRITILGSNLGVQ AGDIQRISVAGRNCSFQERYSVSTRIVCVIEA AETPTTGGVEVDVFGKLGRSPPNVQFTTQQP KFLSVEPQQPQAGATLTHIFGHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMILEVSYGGSPVPNGIFTTYRENPVLRAFE PLRSFASGGRSINVTOGGFSLIGRFAMVVIAEP LQSWQPPREAESLQPMTVVGTTQYFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDFIFENTGGVKKQVNKLJRAR GTNLNKAMTLQBAEAFVGAECTMKTLTET DLYCEPPEVQPPFKRRGKRDTITINLPETIVKF GSREWVLGRVEYDTRYSDVPLSLLBLLVIVWM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMDEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DPEPRRPVVEQALYQFSNLLINSKSFLNFIHIT LVENQPEFSARAKYYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMRRSETVV ERMLSNWMSICLYQYJKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAYTTLNDTGLLGDD VEXPLTINSVIVQDEGVSKRGSVKEKE VKEKIIDQVYRGQPCSCVPRPDSVVLEWPG STAQILSDLDLTSQFEGNWRKNVTLMHYNNX DGATLLISKGVSQQPEDSQQDLJCGERHALL EEERRVWHLVRPTDEVDEGKSKRGSVKEKE		1	1				KHPARNKTLLARMCKEDPNY YSYLEMDLQC							
ACYTGTREADDIFYKPFHGDIQCGGHAPGSGN SFPCGSEHLPYPLGSRDLRGTAVLQRGGIN LTAVTVAAENNHTVAFLGTSDGRILKVYLTP DGTSSEYDSILVENKR VKRULVLSGLIGSLY AMTODKVFRLPVQEGLSYPTCTQCRDSQDPY CGWCVVEGRCTRKAECPRAEEASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLCLGGESPPIPARVEGEAVICNSPSS IPVTPGQDHVAVTIQLLLRRGHTLTSYQYFP YDCRQAMSLEENLPGSCVSNRWTCQWDLR YHECREASPPPEDGIVRAHMEDSCPQFLQPSP LVIPMNHETDVNPQGKNLDTVKGSSLHVOSD LIKFMEPTMOESGTFAFRTFKLSHDANETL PHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGQSRCVYEALCNTT SECPPPVITRIQPETGPLGGGRITILGSNLGVQ AGDIQRISVAGRRCSFQPERYSVSTRIVCVIEA AETPTTGGVEVDVPGKGRSPPNQDFTFQQP KYLSVEPQQPQAGGTTLTHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATR QMLLEVSYGGSPVPNPGIFTTVRENPVLRAFE PLRSFASGGRSINVTQQGFSLQRFAMVVIAEP LQSWQPPREAESLQPHTVVSTQFFAMVIAEP LQSWQPPREAESLQPHTVVSTQFFAMVIAER EAGAFEYVPDPTFFFFTGGVKKQVNKLIRAR GTINLKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRQKRDTTINLPEFIVKF GSREWVLGRVEYDTRYSDVPLSLLPLVIVPM VVVIAVSYCYWKKSQQAEREYEKKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMTTGKL DPEPRRPVVEQALYQFSNLLNSKSFLINFHT LLENQFEFSARAKVYFASLLTVALHGKLEYYT DDMHTLFILLELQYVVAKNPKLMRRSSTVV ERMLSNWMSICLYQYLKDSAGEPLYKIKSQLEG VERMISNNWSICLYQYLKDSAGEPLYKIKSQL VKEKIDQVYRGQPCSCWPRPDSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVVLEWPG STAQILISRUNGSCYPEDGSVSVLEKE DETATLYTITUL ELICEYVALOFYDIPFGSVVL		1			1	1	RDPDIHAAAFGTCLAASVAAPGSGKVLYAVF							
SFPCGSEHLPYPLGSRDGLRGTA/QRGGIN LTAYTVAAENHTV AHGTSDGRILKVYLTP DGTSSEYDSILVEINKRVKRDLVLSGDLGSLY AMTODKVFRLPVQECLSYPTCTOCRDSQDPY CGWCVVCGGCTKRAECPRAEEASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLCLFGESPPHPARVEGEAVICNSPS BYTPPGQDHVAVTIQLLRRGNIFLTSYQYPF YDCRQAMSLEENLPCISCVSNSWTCQWDLR YHECREASPNPEGBIVRAHMEDSCPQFLGSPS LVIPMHETDVNFQGKNLDTVKGSSLHVGSD LLKFMEPVIMQESGTFAFRTFKLSHDANETL PLHLYVKSYGKNIDSKLHVILVDCSFGRSDC SLCRAANPDYRCAWCGGGSRCVYEALCNIT SECPPPYTRIQPETGPLGGRITILGSNLGVQ AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA AETPTTGGVEVDVFGKLGRSPPNVQFTFQQP KYLSVEPQOGPQAGGTTLITHGTHLDTGSQED VRYTLNGVPCKVTKFGAQLQCVTGFQATRG QMLLEVSYGGSPVPNPGIFTTVRENPVLRAFE PLRSFASGGRSINVTQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVYLSPAYDTEPETAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQNKLIRAR GTNLNKAMTLQEAAFVQAGRCTMKTLTET DLYCEPPEVQPPFKRRQKROTTINLPETVKF GSREWVLGRVEYDTRYSDVPLSLLLLVLVTW VYLAYAYYCYWRKSQAEREYEKKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DPEPRRPVVEQALTQYSNLLLNSKSFLNFIHT LYENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMRRSSTVV ERMLSNWMSICLYQYIKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAXYTLNDTGLLGDD VEXPLTLYSVIYQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLLISKYGVSQQPEDSQQDLPGERHALL EEERRVWHLVRPTDEVDEGKSKRGSVKEKE	1	1	1			Į.	SRDSRSSGGPGAGLCLFPLDEVHAKMEANKN							
LTAYTVAAENNHTVALGISDGRILKVYLTP DGTSSEYDSILVENKRVKRDLVLSGDLGSLY AMTQDKVFRLPVQECLSYPTCTQCRDSQDPY CGWCVVEGRCTRKAECPRAEASHWLWSRS KSCVAVTSAQPOMSRRAQGEVQLTVSPLPA LSEEDELLCLFGESPPHPARVEGEAVICNSPSS IPVTPFQQDHVAVTIQLLRRGNIFLTSYQYPF YDCRQAMSLEENLPCISCVSNRWTCQWDLR YHECREASPPEDGIVRAHMEDSCPGFLGPSP LVIPMNHETDVNFQGNLDTVKGSSLHVGSD LLKFMEPVTMQESGTFAFRTPKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAKGGSRCVYEALCNTT SECPPPVITRIQPETGPLGGGRTTLGSNLGVQ AGDIQRISVAGRNCSFQFBEXYSVSTRIVCVIEA AETPFTTGGVEVDYFGKLGRSSPNVQFTFQQP KPLSVEPQQGPQAGGTTLTHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNGGIFTYRENPVLRAFE PLRSFASGGRSINVTGQGFSLQRFAMVVLAEP LQSWQPPRAEASLQPMTVVGTDYYFHDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDFTFENFTGGVKQVNKLURAR GTNLINKAMTLQBEAFAVGAERCTMKTLTET DLYCEPPEVQPPPKRRQKDDTTINLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVIVFM VVVIAVSYYCYWRKSQQAEREYEKKSQLEG LEESVRDRCKKEFTDLMIBEDDQTNDVHEAG IPVLDYKTYTDRVFFIPSKDGGKDVMTTGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFHT LENQPEFSARAKVYTASLLTVALHGKLEYYT DIMHTLFLELLEGDYVASLRTVALHGKLEYYT DIMHTLFLELLEGYVASLLTVALHGKLEYYT DIMHTLFLELLEGYVASLTLVALHGKLEYYT DIMHTLFLELLEGYVASLTLVALHGKLEYYT DIMHTLFLELLEGYTDAVKRYNTLMRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGFVDAVQKAKYTLNDTGLLGDD VEYAPLTVSVTVQDEGVDAPPKVLINCDTISG VKEKIDQVYRGQPCSCWPRDSVVLEWRPG STAQILSDLDLTSQREGGWKKVNTLMTYNVR DGATTLILSKVGVSQOPEDSQOLPGERHALL EEENRVWHLVRFTDEVDGEGKSKRGSVKEKE		İ			}	}	ACYTGTREARDIFYKPFHGDIQCGGHAPGSSK							
DOTSSEYDSILVEINKRYKRIDIVLSGDLGSLY AMTQDKVFRLPVQECLSYPTCTQCRDSQDPY CGWCVVEGRCTRKAECPRAEEASHWLWSRS KSCWAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLCLFGESPHPARVEGEAVICNSPSS IPVTPPQQDHVAVTIQLLLRGNIFLTSYQYPF YDCRQAMSLEENLPCISCVSNRWTCQWDLR YHECRASPNPEDGIVRAHMEDSCPGLGSPS LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD LLKFMEPVTMQESGIFAFRTPKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGQSRCVYEALCNTT SECPPPVTTRIQPETGPLGGGIRITLLGSNLGVQ AGDIQRISVAGRNCSFQFERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGSPNVQFTFQQP KPLSVEPQQGPQAGGTTLTHIGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFTYRENPVLRAFE PLRSFASGRSINVTQQGFSLQRFAMVVIAEP LOSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEFFEAYNLTVLIEMDGHRALLRT EAGAFEVVPDFTFFNGGVKKQVMKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRRQKRDTTHINLPETIVKF GSREWVLGNVEYDTRVSDVPLSLLIPLVTVFM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVADRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVJPSSVPLSLIPLVTVFM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVADRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVJPSSVPLSLIPLVVTHAG IPVLDYKTYTDRVJPSSVPLSLIPLVTVHAG LENGPEFSARAKVYTASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQVLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVTQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLLISKYGVSQQPEDSQQDLPGERHALL EEENRVWHLVRFTDEVDGKSKKRGSVKEKE	İ	İ		1	1	1	SFPCGSEHLPYPLGSRDGLRGTAVLQRGGLN							
DOTSSEYDSILVEINKRYKRIDIVLSGDLGSLY AMTQDKVFRLPVQECLSYPTCTQCRDSQDPY CGWCVVEGRCTRKAECPRAEEASHWLWSRS KSCWAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLCLFGESPHPARVEGEAVICNSPSS IPVTPPQQDHVAVTIQLLLRGNIFLTSYQYPF YDCRQAMSLEENLPCISCVSNRWTCQWDLR YHECRASPNPEDGIVRAHMEDSCPGLGSPS LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD LLKFMEPVTMQESGIFAFRTPKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGQSRCVYEALCNTT SECPPPVTTRIQPETGPLGGGIRITLLGSNLGVQ AGDIQRISVAGRNCSFQFERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGSPNVQFTFQQP KPLSVEPQQGPQAGGTTLTHIGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFTYRENPVLRAFE PLRSFASGRSINVTQQGFSLQRFAMVVIAEP LOSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEFFEAYNLTVLIEMDGHRALLRT EAGAFEVVPDFTFFNGGVKKQVMKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRRQKRDTTHINLPETIVKF GSREWVLGNVEYDTRVSDVPLSLLIPLVTVFM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVADRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVJPSSVPLSLIPLVTVFM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVADRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVJPSSVPLSLIPLVVTHAG IPVLDYKTYTDRVJPSSVPLSLIPLVTVHAG LENGPEFSARAKVYTASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQVLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVTQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLLISKYGVSQQPEDSQQDLPGERHALL EEENRVWHLVRFTDEVDGKSKKRGSVKEKE	1	- I	1			ļ	LTAVTVAAENNHTVAFLGTSDGRILKVYLTP							
AMTODKYFRI.PYOECLSYPTCTOCRDSQDPY CGWCVVEGRCTRIAECFRAEBASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLCLGESPPHPARVEGAVICNSPSS IPYTPPGODHYAVTLQLRGNIFLTSYQYFF YDCRQAMSLEENLPCISCYSNRWTCQWDLR YHECREASPNPEDGIVRAHMEDSCPGFLGPSP LVPMANHETUNYFQGKNLDTVKGSSLHVGSD LLKFMEPYTMQESGTFAFFTFLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGGSRCVYBALCNTT SECPPPYTRIQPETGPLGGGRITILGSNLGVQ AGDIQRISVAGRNCSFQBEXYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP KPLSVEPQQGPQAGGTLTHGTHLDTGSCED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFTYRENPVLRAFE PLRSFASGGRSINVTQGGFSLIQRFAMVVIAEP LQSWQPPRAEASLQPMTVVGTDYVFHNDTK VVFLSPAVEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLJRAR GTNLNKAMTLQEAEAFVGABRCTMKTLTET DLVCEPPEVQPPPKRRQKADTTINLPETIVKF GSREWVLGRVEYDTRVSDVPLSLLDLTVVFM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTINDVHEAG IPVLDYKTYTDRVPLSKDGKDVMITGKL DIPEPRRPVEQALYQFSNLINSKSFLINFIHT LENQPETSARAKVYTASLLTVALHGKLEYYT DIMHTLFLELLEGYVVVAKINFLMLRRSETVV ERMLSNWMSICLYQVLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVTVQDGGVDAIPVKVLNCDTISQ VKEKIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVITLMHTYNKP DGATLILSKVGSQPEDSQQDLPGERHALL EEENRVWHLVRFTDEVDGGKSKRGSVKEKE	1	l.	1		1		DGTSSEYDSILVEINKRVKRDLVLSGDLGSLY							
CGWCVVEGRCTRKAECPRAEEASHWLWSRS KSCVAVTSAQPQNMSRRAQGEVQLTVSPLPA LSEEDELLCLEGESPPHPARVEGEAVICNSPSS IPVTPGQDHVAVTIQLLLRRGNIFLTSYQYFF YDCRQAMSLEENLPCISCVSNRWTCQWDLR YHECREASPNPEDGIVRAHMEDSCPQFLGPSP LVIPMHHETDVNPQGKNLDTVKGSSLHVGSD LLKFMEPVTMQGSGTFAFRTPKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGGSRCVVEALCNTT SECPPPVITTIQPETGPLGGGRITILGSNLGVQ AGDIQRISVAGRNCSPQPERYSVSTRIVCVIEA AETPFTGGVEVDVPGKLGRSPPNVQFFTQQP KPLSVSPQQQPQAGGTTLTHIGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSINVTQGGFSLJQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDFTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFYGAERCTNKTLTET DLYCEPPEVQPPFKRRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLLIPLVTVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG BPVLDYXTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLINSKSFLINFIHT LENGQPEFSARAKVYTASSLTVALHGKLEYYT DIMHTLFLELLEQTVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAT KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSTVTQDEGVBAPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRDSVVLEWRFG STAQILSDLDLTSQREGRWKRNYTLMHYNVR DGATLLISKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE		1		1	1	1	AMTODKVFRLPVOECLSYPTCTQCRDSQDPY							
LSEEDELLCLFGESPPHPARVEGAVICNSPSS IPVTPPGQDHVAVTIQLLLRRGNIFLTSYQYPF YDCRQAMSLEENLPCISCVSNRWTCQWDLR YHECREASPPEDGIVRAHMEDSCPQFLGPSP LVIPMNHETDVNPQGKNLDTVKGSSLHVGSD LLKFMBPVTMQESGTFAFRTPKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGQSRCVYEALCNTT SECPPPVITRIQPETGPLGGGRITTLGSNLGVQ AGDIQRISVAGRNCSPQPERYSVSTRIVCVIEA AETPFTGGVEVDVPGKLGRSPPNVQFTFQQP KPLSVEPQQGPQAGGTTLTHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFTTYRENPVLRAFE PLRSFASGGRSINVTQGGFSLIQRFAMVVIAEP LQSWQPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVQAERCTMKTLTET DLYCCPPEVQPPFKRQKRDTTHNLPEITVKF GSREWVLGRVEYDTRVSDVPLSLLPLVTVPM VVYLAVSVYCYWRKSQAEREYEKIKSQLEG LEESVRDCKKKFTDLMIEMEDQTTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPPERRPVVEQALYQFSNLINSKSFLINFHT LENQPEFSARAVYTASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPYDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCUPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHTYNVR DGATLLISKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE	i	1		1	1	1	CGWCVVEGRCTRKAECPRAEEASHWLWSRS							
LSEEDELLCLFGESPIPHARVEGEAVICNSPSS IPVTPPGQDHVAVTIQLLLRGNIFLTSYQYPF YDCRQAMSLEENLPCISCVSNRWTCQWDLR YHECREASPNPEDGIVRAHMEDSCPQFLGFSP LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD LLKFMEPVTMQESGTFAFRTPKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRANPDYRCAWCGGQSRCVYEALCNTT SECPPYTTRIQPETGPLGGGRITLIGSNLGVQ AGDIQRISVAGRNCSFQERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTFQP KPLSVEPQQGPQAGGTTLTHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFTYRENPVLAFE PLRSFASGGRSINVTQQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPIFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRQKRDTTHNLPETIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYSKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYDRVFFLPSKDGDKDVMITGKL DPEPRRPVVEQALYQFSNLLNSKSFLINFHHT LENQPEFSARAKVYTASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYUKDSAGEPLYKLFKAI KHQVEKGPVDAAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHFINVR DGATLLISKGVSQQPEDSQQDLPGERHALL EEENRVWHLVRFTDEVDEGKSKRGSVKEKE PTX ATTERIJ TRLI SVKRGSVKEKE	1	1	1	1			VSCVAVTSAOPONMSRRAOGEVOLTVSPLPA							
IPVTPPGQDHVAVTIQLLLRRGNIFLTSYQVFY YDCRQAMSLEENLPCISCVSNRWTCQWDLR YHECREASPNPEDGIVRAHMEDSCPQFLGPSP LVIPMHETDVNFQGRNLDTVKGSSLHVGSD LLKFMEPVTMQESGTFAFRTPKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGQSRCVYEALCNTT SECPPPVITRQPETGPLGGGITILGSNLGVQ AGDQRISVAGRNCSFQPERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP KPLSVEPQQGPQAGGTTLITHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSINVTGQGFSLIQRFAMVV1AEP LQSWQPRFAESLQPMTVVGTDYVFINDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTTENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKISQLEG LEESVRDRCKKEFTDLMIEMEDQTINDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPERRPVVEQALYQFSNLLNSKSFLINFHTI LUENQPEFSARAKVYTASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNVMMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLINCDTISQ VKEKIIDQVYRGOPCSCWPRDSVVLEWRPG STAQILSDLDLTSQRGGRWKRNYTLMHYNVR DGATLLISKVGVSQQPEDSQQDLPGERHALL EEERRVWHLVRFTDEVDEGKSKRGSVKKEK			1	1			LEEDELL CLEGESPPHPARVEGEAVICNSPSS							
YDCRQAMSLEENLPCISCVSNRWTCQWDLR YHECREASPNPEDGIVRAHMEDSCPQFLGPSP LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD LLKFMEPVTMQESGTFAFRTPKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRANPDYRCAWCGGQSRCVYEALCNTT SECPPPVITRIQPETGPLGGGRITLIGSNLGVQ AGDIQRISVAGRNCSFQFERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTFQP KPLSVEPQQGPQAGGTTLTHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSNNVTQQGFSLJQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFINDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRRQKROTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVVPM VVVIAVSVYCYWRKSQQAEREYEKKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPPRRPVVEQALYQFSNLINSKSFLINFIHT LUENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSIGLYQYLKDSAGFELYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGGPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRNNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRFTDEVDEGKSKGSVKEKE	Ì	1			1		THE TRUE CONTROL OF THE PROPERTY OF THE PROPER							
YHECREASPNPEDGIVRAHMEDSCPOFLGFSP LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD LLKFMEPPVTMQESGTFAFRTRKLSHDANETL PLHLYVKSYGKNIDSKLHVTLVDCSFGRSDC SLCRAANPDYRCAWCGGQSRCVYEALCNTT SECPPPVTTRIOPETGPLGGGIRTILLGSNLGVQ AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA AETPFTGGVEVDVYGKLGRSPPNVQFTFQP KPLSVEPQQOPQAGGTTLHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFTTYRENPVLRAFE PLRSFASGGRSINVTQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRQKBDTTINLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM VVVIAVSVCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVYCPLSKSGQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFLPSKDGDKDVMTGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFHIT LENOPETSARAKVYTASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIDQVYRQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRTPTDEVDEGKSKRGSVKEKE	1			}			IPVIPOQDITATIQEEDIAGOTI DIOTQ							
LVIPMNHETDVNFQGKNLDTVKGSSLHVGSD LLKFMEPVTMQESGTFAFRTPKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGGSRCVYEALCNTT SECPPPVTRIQPETGPLGGGIRITLGSNLGVQ AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA AETPFTGGVEVDVPGKLGRSPPNVQFTFQQP KPLSVEPQQQPQAGGTTLTHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEVVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPKRRQKRDITHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLLIPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMTTGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT LENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWSICLVQYLKDSAGEPLYKLFKAT KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQLPGERHALL EEENRVWHLVRTPTDEVDEGKSKRGSVKEKE		. [1				YDCKQAMSLEENLFCISCVSNRW TOO WOLK							
LLKFMEPVTMQESGTFAFRTPKLSHDANETL PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAAAPPDYRCAWCGGQSRCVYEALCNTT SECPPPVTTRIQPETGPLGGGIRITILGSNLGVQ AGDIQRISVAGRNCSFQFERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP KYLSVSPQQGPQAGGTTLTHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSINVTQGGFSLIQRFAMVVIAEP LOSWQPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEFSVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLINSKSFLINFHT LENQPEFSARAKVYTASLLTVALHGKLEYYT DINHTLFLELEGVYVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKKKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRFTDEVVDEGKSKRGSVKEKE	1			1			YHECKEASPNPEDGIVKAHMEDSCPQFEDISI							
PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC SLCRAANPDYRCAWCGGQSRCVYEALCNIT SECPPYITRIOPETGPLGGGRITITLGSNLGVQ AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP KPLSVEPQQGPQAGGTTLTHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCFRRQKKQVTNKLIRAR GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT LIENQPEFSARAKVYTASLLTVALHGKLEYYT DINHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGGDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRFG STAQUISDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE	1	1	1		1.	ł	LVIPMNHETDVNFQGKNEDTVKGSSLHVGSD							
SLCRAANPDYRCAWCGGQSRCVYEALCNTT SECPPPVITRIQPETGPLGGGIRITLGSNLGVQ AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP KPLSVEPQQGPQAGGTTLTIHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVVVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESYNDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLINSKSFLINFIHT LENQPEFSARAKVYTASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGOPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE PTK ATTENT TRIL SVKGTLOOFVDNFFQSVL		l.	1		İ		LLKFMEPVTMQESGTFAFKTPKLSHDANETL							
SECPPVITRIQPETGPLGGGRITLGSNLGVQ AGDIQRISVAGRNCSFQPERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP KPLSVEPQQQFQAGGTTLTHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PPLRSFASGGRSINVTQGGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRYSDVPLSLLIPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLNFIHT LVENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VEKKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEERNVWMLVRPTDEVDEGKSKRGSVKEKE PTYATTETYLTRILSVKGTLOOFVDNFFOSVL	-	1	Į	1		1	PLHLYVKSYGKNIDSKLHVTLYDCSFGRSDC							
AGDIQRISVAGRNCSFQFERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP KPLSVEPQQGPQAGGTTLTIHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFYGAERCTMKTLTET DLYCEPPEVQPPPKRRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVTVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT LVENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEERNVWMLVRPTIDEVDEGKSKRGSVKEKE PTYATTERYLTRILSVKGTLOOFVDNFFGSVL		1			1		SLCRAANPDYRCAWCGGQSRCVYEALCNII							
AGDIQRISVAGRNCSFQFERYSVSTRIVCVIEA AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP KPLSVEPQQGPQAGGTTLTIHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFYGAERCTMKTLTET DLYCEPPEVQPPPKRRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVTVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT LVENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEERNVWMLVRPTIDEVDEGKSKRGSVKEKE PTYATTERYLTRILSVKGTLOOFVDNFFGSVL)		1	ļ	ł	1	SECPPPVITRIQPETGPLGGGIRITILGSNLGVQ							
AETPFTGGVEVDVFGKLGRSPPNVOFTFQP KPLSVEPQQGPQAGGTTLTIHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT LENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE PTYLATIFIJI TRILSVKGTLOOFVDNFFOSVL	1	ì	- 1	1		Ì	AGDIORISVAGRNCSFQPERYSVSTRIVCVIEA							
KPLSVEPQQGPQAGGTTLTHGTHLDTGSQED VRVTLNGVPCKVTKFGAQLQCVTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSINVTQGGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTTENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRRQKRDTTHINLPETIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLINSKSFLINFHT LENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE PUX ATTEIVLTBLL SVKGTLOOFVDNFFOSVL			- 1	1	ł		AETPFTGGVEVDVFGKLGRSPPNVQFTFQQP							
VRVTLNGVPCKVTKFGAQLQCYTGPQATRG QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPFKRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRYSDVPLSLLIPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT LVENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE		 	- 1		Į	1	KPLSVEPOOGPOAGGTTLTIHGTHLDTGSQED							
QMLLEVSYGGSPVPNPGIFFTYRENPVLRAFE PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTINDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT L\ENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE		l	- 1	ļ	ĺ		VRVTLNGVPCKVTKFGAOLOCVTGPQATRG							
PLRSFASGGRSINVTGQGFSLIQRFAMVVIAEP LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT L\ENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE	1			1	ì		OMI LEVSYGGSPVPNPGIFFTYRENPVLRAFE							
LQSWQPPREAESLQPMTVVGTDYVFHNDTK VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTIGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT LIENQPEFSARAKVYTASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE	-	ł	1		1		DI DEFA GGESTIVITGOGEST TORFAMVVIAEP							
VVFLSPAVPEEPEAYNLTVLIEMDGHRALLRT EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVTVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT LIENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLFGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE	ļ		İ	i	I	ì	LOCKORDE A EST OPMTVVGTDVVFHNDTK							
EAGAFEYVPDPTFENFTGGVKKQVNKLIRAR GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT LVENQPEFSARAKVYTASLLTVALHGKLEYYT DINHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE		l	1		1	1	LUBWOFFREADSON INTO TOTAL THE AND CHRAITRT							
GTNLNKAMTLQEAEAFVGAERCTMKTLTET DLYCEPPEVQPPPKRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRYSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT LVENQPEFSARAKVYFASILTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE		1		i	1	i	A ALPANALEELEY I MET A FIRM POLICY FOR							
DLYCEPPEVQPPPKRRQKRDTTHNLPEFIVKF GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT LVENQPEFSARAKVYFASILTVALHGKLEYYT DIMHTLFILELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE	-		ı	1	ļ	1	EAGAPET VPDP I TENT I GOVERNMENT TET							
GSREWVLGRVEYDTRVSDVPLSLILPLVTVPM VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPPRRPVVEQALYQFSNLLNSKSFLINFIHT LYENQPEFSARAKVYTASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE	1		1			J	GINLNKAMILQEAEAFVGAEKCIMKILIEI							
VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTINDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT L\ENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE			- {	-{			DLYCEPPEVQPPPKRRQKRDTTHNLPEFTVKF							
VVVIAVSVYCYWRKSQQAEREYEKIKSQLEG LEESVRDRCKKEFTDLMIEMEDQTINDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT L\ENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE	1				ļ		GSREWVLGRVEYDTRVSDVPLSLILPLVIVPM							
LEESVRDRCKKEFTDLMIEMEDQTNDVHEAG IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT L\ENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE			1	- 1		1	VVVIAVSVYCYWRKSOOAEREYEKIKSQLEG							
IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT LIENQPEFSARAKVYTASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE	1		1 .	-	1		I FESUR DRCKKEFT DLMIEMEDOTNOVHEAG							
DIPEPRRPVVEQALYQFSNLLNSKSFLINFIHT LVENQPEFSARAKVYFASILTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE		1		ı			IPVLDYKTYTDRVFFLPSKDGDKDVMITGKL							
L\ENQPEFSARAKVYFASLLTVALHGKLEYYT DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE PTV ATTEIVLTRIL SVKGTLOOFVDNFFQSVL	Ī	- 1	1	1			DIPEPRRPVVEOALYOFSNLLNSKSFLINFIHT							
DIMHTLFLELLEQYVVAKNPKLMLRRSETVV ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYPVAL DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE PTV ATTEVYLTRIL SVKGTLOOFVDNFFQSVL		- [1			1	LIENOPEESARAKVYFASLLTVALHGKLEYYT							
ERMLSNWMSICLYQYLKDSAGEPLYKLFKAI KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE PTWAITEIYLTRILSVKGTLOOFVDNFFOSVL	- 1	1	-	1	1	1	DIMUTE EL EL EOVVVAKNPKI MI RRSETVV							
KHQVEKGPVDAVQKKAKYTLNDTGLLGDD VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE PTWAITEIVLTRILSVKGTLOOFVDNFFOSVL	1	1	ļ	1	1		EDMI SMIMSICI VOVI KDSAGEPI VKI EKAI							
VEYAPLTVSVIVQDEGVDAIPVKVLNCDTISQ VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE PTWAITEIVLTRILSVKGTLOOFVDNFFOSVL	1	İ		i	1	1 .	EXMEDITE OF A MONTH A FORT A STATE OF THE ST							
VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE PTWAITEIVLTRILSVKGTLOOFVDNFFOSVL	-	1	1	1	- i		KHQVEKGrvDAVQAAAA I ILIVD I GEEGDD							
STAQILSDLDLTSQREGRWKRVNTLMHYNVR DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE PTKATTETYLTRILSVKGTLOOFVDNFFQSVL		1	- 1	l l	1	1	VEAVELA2211A0DECADALLA VALUEDI 120							
DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE PTV ATTETYL TRI L SVKGTLOOFVDNFFQSVL		1	-		İ		VKEKIIDQVYRGQPCSCWPRPDSVVLEWRPG							
DGATLILSKVGVSQQPEDSQQDLPGERHALL EEENRVWHLVRPTDEVDEGKSKRGSVKEKE PTV ATTETYL TRI L SVKGTLOOFVDNFFQSVL	1	- 1	1	İ	ł		STAQILSDLDLTSQREGRWKRVNTLMHYNVR							
EEENRVWHLVRPTDEVDEGKSKRGSVKEKE PTK ALTERYL TRI L SVKGTLOOFVDNFFOSVL	-		1	- 1		1	DGATLILSKYGVSOOPEDSOODLPGERHALL							
		1	- }	ļ		}	EEENRVWHLVRPTDEVDEGKSKRGSVKEKE							
APGHAVPPAVKYFFDFLDEQAEKHNIQDEDTI	1		- [l l	DTW AFTERVI TRLI SVKGTLOOFVDNFFQSVL							
	1		1	Ì	- 1	1.	APGHAVPPAVKYFFDFLDEQAEKHNIQDEDTI							
							1							

NO. of No	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence coide seq- correspondi location (9496 purche where the corresponding to the correspondi					1		Amino acid sequence (A=Atanine C=Cysteine,
Sociation Soci			1100	1			E-Dhearlainine C-Chaine Hall will
sequence with the corresponding a fig. first, and in seal direction and residue of peptide sequence peptide sequence peptide sequence peptide sequence per company to the company of the c			1				r=rnenymannne, G=Glycine, H=Histidine,
uenice aliano acid residue of peptide sequence peptide se	1		ſ				1=Isoleucine, K=Lysine, L=Leucine,
### squence sq	1 -	uence	I				
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Siop codon, *possible uncloulde deletion, *possible uncloude deletion, *possible uncloude deletion, *possible uncloude deletion, *possible uncloude deletion, *possible uncloude deletion, *possible uncloude deletion, *possible uncloude deletion, *possible uncloude deletion, *possible uncloude deletion, *possible unclouded deletion, *possi	uence	1	ļ	914			
peptide squence		1	1		L		
			ł			sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
HIWKTNSLPLRFWYNLIKNPHIFIDVHYHEV	1	i	ľ	ľ		l	/=possible nucleotide deletion, \=possible
BASIS VIAQITMOACTRITERIUS SRDSPSNIKLI YAKEISTYKKIMEN YYKGIRQMYOYSODOM NTHIAEISRAHTDSINTLVALHQLYQYTQKY YDEINALEEDPAQKMQLARRQQIAAALE NKVTDI 839 2189 A 6872 1 1485 RARRILAIQCHYCVCALTDGEQGGRILDGGT WINFSCYCPSLODNSYSSTYTYECDEDPVSLH EQTIOSSURDENNKENPYDAGAL VIEHAPP SWEPQQQNVSATYLUDSVLRSPMGMYESRIKY KSIFKABCSRHGESGDEHVYSSOSCOVRA GTPAHESPQNNAFKOQETIVALQPRIDORTAT SPKDAFTRYQDLANEEAAQVHGYKDPAPAS TOSVLANDGTDSADPSPYHRDGQNISCAGNAF GTPAHESPQNNAFKOQETIVALQPRIDORTAT SPKDAFTRYQDLANEEAAQVHGYKDPAPAS TOSVLANDGTDSADPSPYHRDGQNISCAGNAF GOKKYRSHODKTSNIPSVLKWANDIAACKR QUKELKILSSEQGSAPKGPFRNILCEQPTVP RENOKFRAGFPESSGETFOAALTCHKER EQLPPQEDSKYTKQDKNILKHYDRYRIKQIL STPILITIVSQDTCALLLCTDV RENOKFRAGFPESSGETFOAALTCHKER EQLPPQEDSKYTKQDKNILKHYDRYRIKGIL STPILITIVSQDTCALLCTDV SRSCCSGSLSSYDYSEDFLCDCSEKAINRN YLQQVVKEKEKKYNYNSKIGSKGKGEISV EKKHTWNASLFNSQHMAQRDAMAERILS AALHKIKGIKKELADMHHIKLEALITENOFLK QUQLRHLKAIGKYENSQNNLPQMAEROHE KNLRQLLRKSGEKERTLSKRLETIDSQLLET KNLAQTATKTLQVEVKHLQQKLKEKDREL EENIYSPRILLENTHOTYNKENSCHALBER EENIYSPRILLENTHOTYNKENSCHALBER EENIYSPRILLENTHOTYNKENSCHALBER EENIYSPRILLENTHOTYNKENSCHALBER EENIYSPRILLENTHOTYNKENSCHALBER EENIYSPRILLENTHOTYNKENSCHALBER SKYLKYNSKISCHALBER SKYLKYKKYSLOVICHTON SKYKY BDLSGESKHLEVQILLENTGRQDLKT KNDIAGALGKSEBKRITCHSVLLETT KTLAAQTATKTLQVEVKHLQQKLKEKDREL EENIYSPRILLENTHOTYNKENSCHALBER EENIYSPRILLENTHOTYNKENSCHALBER EENIYSPRILLENTHOTYNKENSCHALBER SKYLKYKYSLOVICHTORKSKYENDEDVILLET NDHKEKSTEDHIEDPTYNVSSTKSWQAD RKILPTTSMRHQOTQKSDVPPLTTKGKKATD NDHKEKSTEDHIEDPTYNVSSTKSWQAD RKILPTTSMRHQOTQKSDVPPLTTKGKKATD OSSPGVAKGSEFLQSKESHPLPSQASTSHA FODSKYTVINSKEPSTEGKKIII 841 2191 A 6874 3 2867 STREMEERESLERGQILLDSGATSHA FODSKYTVINSKEPSTEGKKIII 842 SERGER SKREMERESCHALDSOOYEPSTGKS SKRYKYKKASSASSAGAQKSCHEFF DTPWSDQRRREGGEPROQUQRSPTRARG TCSVEDPYTVISGRASSASSAGAQKSCHEFF DTPWSDQRREGGEPROQUQRSPTRARG TCSVEDPYTVISGRASSASSAGAQKSCHEFF DTPWSDQRREGGERPROQUGRSPTRARG TCSVEDPYLLCQCEFFORKNYKKYSLOVALDR VKASSAGAMANKYEKQULADPEKPRKNATS SKRYSSASSASSAGAGARSCONGES SKORGRADASAGSPRQABASSDVTCKTNAS SKRYSSASSASSAGAGARSCONGES SKRYKYKKASSASSASSAGAGARSCONGE					sequence		nucleotide insertion
YAKEISTYKKIM/ED/YYKGIRQM/QVSDQDM NHIALEISRAHITISIAINI/VAILQI/YYTOKY YDEINALEEDPAAQKMQLARILQQIAAALE NKVTUL 839 2189 A 6872 1 1485	1		İ				HIWKTNSLPLRFWVNILKNPHFIFDVHVHEVV
YAKEISTYKKIM/ED/YYKGIRQM/QVSDQDM NHIALEISRAHITISIAINI/VAILQI/YYTOKY YDEINALEEDPAAQKMQLARILQQIAAALE NKVTUL 839 2189 A 6872 1 1485		1					DASLSVIAQTFMDACTRTEHKLSRDSPSNKLL
MTHLAEISRAHIDSLATIL/ALHQLYQYTQKY YDEINALEEDPAAQKMQLARRQQIAAALE NAVTOL	1	1	1			1	YAKEISTYKKMVEDYYKGIROMVOVSDODM
S39 2189 A 6872 1 1485 RARRILAUCHVCVCALTPGEOGGRILEGOT WAMPSCCPSI, ODNASTITYTECLOPYSLH EDOTDCSSLRDENVKBNYPDAGALVEHAPP SWEPQQNYATVL/DSVLRPSMORTESKRY SUSPACIONAL SERGESHEES SWEPQQNYATVL/DSVLRPSMORTESKRY SUSPACIONAL SERGESHEES SWEPQQNYATVL/DSVLRPSMORTESKRY SUSPACIONAL SERGESHEES SWEPQQNYATVL/DSVLRPSMORTESKRY SUSPACIONAL SERGESHEES SWEPQQNYATVL/DSVLRPSMORTESKRY SWEPQQNYATVL/DSVLRPSMORTESKRY SWEPQQNYATVL/DSVLRPSMORTESKRY SWEPQQNYATVL/DSVLRPSMORTESKRY SWEPQQNYATVL/DSVLRPSMORTESKRY SWEPQQNYATVL/DSVLRPSMORTESKRY SWEPQQNYATVL/DSVLRPSMORTESKRY SWEPQQNYATVL/DSVLRPSMORTESKRY SWEPQQNYATVL/DSVLRPSMORTESKRY SWEPQQNYATVL/DSVLRPSMORTESKRY SWEPQQNYATVL/DSVLRPSMORTESKRY SWEPQQNYATVL/DSVLRPSMORTESKRY SWEPQQNYATVL/DSVLRPSMORTESKRY SWEPQQNYATVL/SWEPQCK/SWEPANLACLEOPTVP RENGKPEAAGPPSSSGETPDAALTCL/KERR SOLPPODSKYTKQDKNIKPLYDRYRIIKQL STP3LIPTIVSQDTCALLLCTDV FRRYTSPIRILFAMSLADILTKTNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKTNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKTNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKTNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKTNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKTNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKTNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKTNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKTNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKTNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNITTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNITTKNIDEHFFGVAL FRRYTSPIRILFAMSLADILTKNITTKNITTKNITTKNITTKNITTKNITTKNITTKN						1	NTHLAEISRAHTDSLNTLVALHOLYOYTOKY
Say 2189 A 6872 1 1485 RARRIALQCHVCVCALTPGEQSGRILPGQT WLMFSCFCPSLQDNSFSSTVTECDEDPVSLP EDQTTOCSSLADENKENYPDAGALVEEHAP EVGTTOCSSLADENKENYPDAGALVEEHAP EVGTTOCSSLADENKENYPDAGALVEEHAP SWEPQQONVEATULVDSVLRPSMGNFKSRRV KSIFKAESGRSHGESQRTEHVVSSQSCFGALVEEHAP SWEPQQONVEATULVDSVLRPSMGNFKSRRV KSIFKAESGRSHGESQRTEHVVSSQSCFGAALVEEHAPA SPKDAFETRQDLNEEEAAQVIGVUCDPAPAS TQSVLAUGITDSADPSVHKDQDREADSAPE DLHSVGTSRILLLYHITDGDNFTAVRRIGGSLF SQSQGRNILDPSAPSPSTQCMPMRSSSRC SCGDOKEPQITTQLTKHIGSLKRKIKFEFEKF QEKKYRPSHGDKTSNPEVLKWMDLAKGRK QUKELKLKLSEEQGSAPKGPRNILCEQPTVP RENGKPEAAGPFPSSGETPDAALTCLKERR EQLPPQEDSKVTKQDKNLKPLYDRYRIIKQIL STPSLIPTIVSQDTCMLLICTDV RENGKPEAAGPFPSSGETPDAALTCLKERR EQLPPQEDSKVTKQDKNLKPLYDRYRIIKQIL STPSLIPTIVSQDTCMLLICTDV STPSLIPTIVSQDTCMLTATE STPSLIPTIVSQDTCMLTATE STPSLIPTIVSQDTCMLTATE STPSLIPTIVSTATIVSQDTCMLTATE STPSLIPTIVSQDTCMLTATE STPSLIPTIVSQDTCMLTATE STPSLIPTIVSTATIVSQDTCMLTATE STRIPTIVSTATIVSQDTCMLTATE STRIPTIVSTATIVSQDTCMLTATE STRIPTIVSTATIVSQDTCMLTATE STRIPTIVSTATIVSQDTCMLTATE STRIPTIVSGDTCMLTATE STRIPTIVSTATIVSQDTCMLTATE STRIPTIVSGDTCMLTATE STRIPTIVSTATIVSQDTCMLTATE STRIPTIVSGTCMLTATE STRIPTIVSGTCMLTATE STRIPTIVSGTCMLTATE STRIPTIVSGTCMLTATE STRIPTIVSGTCMLTATE STRIPTIVSGTCMLTATE STRIPTIVSGTCMLTATE STRIPTIVSGTCMLTATE STRIPTIVSGTCMLTATE STRIPTIVSGTCMLTATE STRIPTIVSGTCMLTATE STRIPTIVSGTCMLTATE STRIPTIVSGTCMLTATE STRIPT		i					YDEINALEEDPAAOKMOLAFRLOOIAAALE
839 2189 A 6872 1 1485 RARRIALOCHYCVCALTPGEGGGRELPGOT WLAPSCFCCELQDNSSSTYTYCEDEVSLH EDQTDCSSLRDENNKENYPDAGALVEEHAPP SWEPQQNVEATVLVDSVLRPSMONFKSKYP KSIKKABGGREGGEGTEHVYSSSSCCVRA GTPAHESSQNNAFKQETVRLNQPRIDQRTAT SKDAPFETRQIDABEAQAVHCHVDRPAPAS TQSVLANDGTDSADPSVYHKDGQNEADSAPE DLHSWGTSRLLLYHTTDGDNFTAVRIGCSLF SGGSGRFNIDPESAPSPSTQOFMMPRSSST COKERCE OKEKYRPSHGDKTSNPEVLKWANDLAKGENG (LKELKLKLSEEGGSAPKGPRFNLCEQPTVP RENGKPEAAGPEPSSSGGETPDAALTCLKERR EQLPPQEDSKVTKQNSKHIKPLYDBYRTIKQIL STPSLIPTTVSQDTCMLLICTDV RENGKPEAAGPEPSSSGGFTPDAALTCLKERR EQLPPQEDSKVTKQNSKHIKPLYDBYRTIKQIL STPSLIPTTVSQDTCMLLICTDV RENGKPEAAGPEPSSSGGFTPDAALTCLKERR EQLPPQEDSKVTKQNSKHIKPLYDBYRTIKQIL STPSLIPTTVSQDTCMLLICTDV SSSCCSGSLSSQVDYSEDPLCDCSEKAINRN YLKQVVKEEKKKKYNVSKISQKGEISV EKKHTWNASLPNSQHMIAQRRDAMAHRILL ARLHKIKGLKNELADMHHIKLALILTENOFILK QLQLRHLKAIGKYENSQNNLPQIMAKHQNEV KNLRQLLRSGDKERT SRKLRAFDOLLKT KDILQALQKLSBDKNLAREFELTHKLSHITTK MDAANDKKIQSLEKERTI SRKLRAFDOLLKT KDILQALQKLSBDKNLAREFELTHKLSHITTK MDANDKKIQSLEKERTI SRKLRAFDOLLKT KDILQALQKLSBDKNLAREFELTHKLSHITTK MDANDKKIQSLEKERTI SRKLRAFDOLLKT KDILQALQKLSBDKNLAREFELTHKLSHITTK MDANDKKIQSLEKERTI SRKLRAFDYRK STKSVQAD RICHFTSMRHIGOTKSDVPPLITKGKKATOL NIDHKEKSTEINHEIPHCVNKLPKQEDSKRYF EDLSGEEKHLEVQULLENTGQKDKKEDQEK KNIFVKEEQGLPKRITTKGKKATOL NIDHKEKSTEINHEIPHCVNKLPKQEDSKRYF EDLSGEEKHLEVQLLENTGGKDKKEDQEK KNIFVKEEQGLPKRITTKGKKATOL NIDHKEKSTEINHEIPHCVNKLPKQEDSKRYF ERKFSMQRNGVDTULGKGTAPTSGGYVLKTU QSSGVAKGSEPL QSSGVAKGSEPL GORGVAGGSEPL GORGVAKGSEPL GORGVAKGSEPL GORGVALSEPL GORGVAGGSEPL GORGVACAGGSEPLOT GORGVAGGSEPL GORGVAGGSEPL GORGVAGG	1	J				J	
### WILMFSCFCSLQDNSFSSTTVTECDEDPYSIL EDOTTOSSLADEDNIKENTPDAGALVEEHAPP SWEPQQONVEATULVDSVLRPSMGNRESRIKP KSIFKABSGRSHCSQQTEHVVSSQSCQVRA GTPAHESPQNNAFKCQETVRLIQPRIDQRTAT SPKDAFETRQDLNEEEAAQVHGVKDPAPAS TQSVLANDGTDSADPSPVHKDQDRDSAPE DLHSVGTSRLIL/YHITDGDNFTAVRHGCSLJ SQSQSGRNLDPSAPSPSTQCPMMPRSSSRC SCGDGKEPQITTQLTKHIGSLKRKIRKFEEKFE QEKKYRPSHGDKTSNPEVLKWANDLAKGRK QLKELKLKLSEEQGSAPKGPRNLLCEQPTVP RENGREAAGPEPSSGGETPDAALTCLKERR EQLPPQEDSKVTKQDKNLIKPLYDRYRHKQIL STPSLIPTIVSQDTCMLLLCTDV FREYFSTRLIAMSLADLTKTNIDEHFFGVAL ENNRSSAACKRSPGTGDFSRNSNASNKSVDY SRSQCSGSLSSQVDYSEPFLCDESKANTN YLKQPVVKEKEKKKYNVSKISQSKGGKEISV EKKHTWNASLFNSQNIHMAQRRDAARINLS ARLHKIGLKNELADMHHLEALITENOFILK QLQLARLKLAIGYVSNQNLIKPLANETHSQLKT KDILQALQKLSEDKNLAREFELTHLSIITTK MDANDKKIQSLEKQLRICKARSRQLAETR KTLAAQTATKTLQVEVKHLQKLKEKDERJ SMRTHSTFLAKHLDTEDYPKVSSKSVQAD RRILPFTSMRHGGTQKSDVPPLTTKGKKATI NDIKKESTEINHEIPHCVNLLYKEDSKRYY EDLSGEEKHLEVQILLENTGRQKDKKEDQUK KNIFVKEGQELPFKIEVHPERSSNOEDVLVR EKFKRSMQRNGVDDTLGKGTAPYTKGPLRQ RRHTSFTEATENLHHIGLPASGGPANAGNMR YSHSTGKHLSNREEMLEHISDSYPSORDEDVLVR EKFKRSMQRNGVDDTLGKGTAPYTKGPLRQ RRHTSFTEATENLHHIGLPASGGPANAGNMR YSHSTGKHLSNREEMLEHISDSYPSORDEDVLVR EKFKRSMQRNGVDDTLGKGTAPYTKGPLRQ RRHTSFTEATENLHHIGLPASGGPANAGNMR YSHSTGKHLSNREEMLEHISDSYPPSORTSHA FGDSKVTVVNSIKPSSTFEGKRUII SSRTGKSSHGPSWRKKYSLVNRPPGSSDPPA DHAVRPIHGARGQOPPVQGPVLERQVQLS QQQNVVKKYPFSKSGSASASGAQRGSLEETF DTPWSDQRFEGGEGFPRGQLLDGPRSTRARG TCSVEDPLLVCQKEPGFPRRVKSVGSVOGDS REPRRTYSESTIAVKASPSSABPSRFTOVALG RRLGSHSVASCAPQLLGDRRVDAGHTDQVPS GGSVGGARPASGFPRQAREASLVVTCKTNRP RKNDYKWAASSKSSPRVARRALSPRVAARI VCKASAGMANKVEKPQLLADPEKPRPATS SKPGASRSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSKSSPGDRPALAHSIGKKSI GETELSAYKVKTKTIKBRRGSTSLFPGOKKSI GETELSAYKVKTKTIKBRRGSTSTLFDGKKSI GETELSAYKVKTKTIKBRRGSTSTLFDGKKSI GETELSAYKVKTKTIKBRRGSTSTLFDGKKSI GETELSAYKVKTIKTIKBRRGSTSTLFDGKKSI GETELSAYKVKTKTIKBRRGSTSTLFDGKKSI GETELSAYKVKTKTIKBRRGSTSTLFDGKKSI GETELSAYKVKTKTIKBRRGSTSTLFDGKKSI GETELSAYKVKTIKTIKBRRGSTSTLFDGKKSI GETELSAYKVKTIKTIKBRRGSTSTLFDGKKSI GETELSAYKVKTIKTIKBRRGSTSTLFDGKKSIG GETELSAYKVKTIKTIKRRGSTSTLFDGKK	839	2189	A	6872	1	1485	
EDOTDCSSLADENNKENTPDGALVEEH, SWEPAGOQNVEATV. VDSVLRPSMONEKSRKP SWEPAGOQNVEATV. VDSVLRPSMONEKSRKP KSIKABSGRSHGESQETEHVVSSQSECOWA GTPAHESSPONNAFKCQETIVRLOPRIDQRTAT SPKDAFETR\QDLNEEBAQVHIOVCDPAPAS T\QSVLAUGTDSADPSVVIKDQQNBADSAPE DLHSV\GTSRLLLVHTDGDNPTAVRHGCSLF SQQSQRNLDPESAPSPSTQFMMPRSSSR SGGDGKEPQTITICITKHIGKJKKRKFKEEKFE QEKKYRPSHGDKTSNPEVLKWANDLAKGR QLKELKLKLSEEGGSAPKGPFRNLCEQPTVP RENGKPEAAGPEPSSSGEETPDAALTCLEERR QLXPPGEDSSKVTKQDKNLIKLYPSKIKGTEEKFE QEKKYRPSHGDKTSNPEVLKWANDLAKGR QLKELKLKLSEEGGSAPKGPFRNLCEQPTVP RENGKPEAAGPEPSSSGEETPDAALTCLEERR EQLPPGEDSSKVTKQDKNLIKLYPSKIKGL STPSLIPTHVSQDTCMLLLCTDV FREYFSFIRLFAMSLADLTKNIDEHFFGVAL ENNRSAACKRS*GTGDPSINNSASNKSVDY SRSQCSGSGSSQVTYSEDFLCDCSEKANNN YLKQVVKEEKKKKYNVSKISQSKGQKEISV EKKHTWNASLFNSQHMMQRRDAMAHRILL ARLHKIRGLNELALDMHHILE ALLHENOFLK QLQLIKHLKAIGKYENSQNNLPQMAKHQUKK KNLRQLLKSSGEKERTLSRKLRETTSQLLKT KDILQALQKLSEBOKALAEREELTHLSINTTK MDANDKKIQSLEKERTLSRKLRETDSQLKT KDILQALQKLSEBOKALAEREELTHLSINTTK MDANDKKIQSLEKGERTLSRKLRETDSQLKT KDILQALQKLSEBOKALAEREELTHLSINTTK MDANDKKIQSLEKGETLSRKLRETDSQLKT KDILQALQKLSEBOKALAEREELTHLSINTTK MDANDKKIQSLEKGETLSRKLRETDSQLKET KNLRVSHELKALBUTEDDYKVSSTKSVQAD RKLLPFTSMRHQGTQKSSONLPPQMAKHQUK EEKKRSNQRNGVDDTULKGKTAPTKGFLAR EEKSKSNQRNGVDDTULKGKTAPTKGFLAR VSHSTEATHHIGLPASGGPANAGMR SKITTSTRENGENGENCTV REKERSNQRNGVDDTULKGKTAPTKGFLAR RHYSFTEATENLHHGLPASGGPANAGMR SKITTSTREMEEKELHSDSQVPSFSGKS SIKVNDTTFRDKKSSLMELFGSGYVLKTD QSSPGVAKGSEEPLGSKSSHLPPSQATSSHA FODSKVTVVNSIKPSSPTEGGRKHI 841 2191 A 6874 3 2867 SSRTREMEEKELHSDSQVPSFSGKS SIKVNDTTFRDKKSSLMELFGSGYVLKTD QSSPGVAKGSEEPLGSCSPHLPVLREQVQLS QQNVVKKPPSKSGSASASQAQRGSLEFE DTPWSDQKPREGGEPPRGULQGRSPTPAGLPSTRAFE TCSVEDPLLVCQKEPGRPRRVKSVGSVQDS REPRRTYVSSVIAVKASPSSASSSSSRYNGSEAG SKDHASQLSPVLSKSSPTQASASAYVARALSPRVAAEN VCKASAGMANKVEKPQLLADPEPKPRATATS SKPQSAPSKYKWKASSPSASSSSSRYRWQSEAG SKDHASQLSPVLSKSSPGGRAPAATSPCATATS SKPQSAPSKYKWKASSPSASSSSSRSFWQSEAG SKDHASQLSPVLSKSSPGGRAPAATSPCATATS SKPQSAPSKYKWASSSPSASSSSSFRWQSEAG SKDHASQLSPVLSKSSSGGGRPAAGHAGALSPRVAAEN VCKASAGMANVEKPQLLADPEPKPRATATS SKPQSAPSKYKWASSSSSSGROBAGAGAGAG	1				-		
SWEPQQONVEATVLVDSVLRPSMGNPKSIKP KSIKRAESGRSHGESQEET AGTAHESPQNNAFKCQETVKL\QPRIDOPTAT SPKDAFETRIODINEELAQVIGNUDPAYAS TQSVLA\DGTDSADPSPVHKDQQNEADSAPE DLHSVGTSRLLLYHTIDGDNPTAVRHGGLIF SGGQRFNLDESASPPSTQOPMMPRSSSRC SCCDGREPQTITQLTKHQSLKRKIRKFEEKFE QEKKYPRSHGDKTSNPSVLKMDLAKGRK QLKELKLKLSEQGSAPKGPPRNLLCEDPTVP RENGKPEAAGPPSSSGEETPDAALTCLKERR EQLPPQEDSKVTKQDKNLIKPLYDRYRIIKQIL STPSLJETIVSQDTCMLLLCTDV RENGKPEAAGPPSSSGEETPDAALTCLKERR EQLPPQEDSKVTKQDKNLIKPLYDRYRIIKQIL STPSLJETIVSQDTCMLLLCTDV SRSQCSCGSLSSSQTVDSEDFLOSEKANRN YLKQPVVKEKEKKXVNVSKISQSKGQKEISV KKHTWNASLFNSQTHMLAGNAMHRILL ARLHKIKGLKNELADMHKLEALLTENOFLK QLQLRIIKANGKYENSQNLOCSEKANRN YLKQPVVKEKEKKXVNVSKISQSKGQKEISV KKHTWNASLFNSQTHMLAGNAMHRILL ARLHKIKGLKNELADMHKLEALLTENOFLK QLQLRIIKANGKYENSQNLOCSEKANRN MDANDKKIQSLEKQLRINCRAFSRQLAETR KTLAAQTATKTLQVVKHLQQKLKEKDRL ERNIYSHILKNLHDTEDYTKSTSSVQAM RKLIPFTSMRHQGTQKSDVPPLITTIKGKATO NIDHKEKSTEINHEIPHCVNKLFKQEDSKRKATO NIDHKEKSTEINHEIPHCVNKLPKGEDSKRKATO NIDHKEKSTEINHEIPHCVNKLPKGEDSKRKATO NIDHKEKSTEINHEIPHCVNKLPKGEDSKRKY EDLSGEEKHLEVQILLEHTGRQKDKKEATO NIDHKEKSTEINHEIPHCVNKLPKGEDSKRKTU REFKESMQRNGVDDTLGGGTAPYTKGPLRO REFKESMQRNGVDDTLGGGTAPATAGNAFA RYSHSTGALLKNLDTLCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1	ļ				İ	FDOTDCSSI RDENNKENVPDAGAI VEEHADD
KSIFKAESGRSHGESOETEHVVSSQSECQVRA GITPAHESPONNAFKCGETVSQPEIDQRTAT SPKDAFETRIQDLNEEEAAQVHGVKDPAPAS TQSVLAUGTDSADPSPVHKDGQNEADSAPE DLHSVGTSRLLLYHTIDGDNTPAVRHGCSJF SGGSQRFNLDPESAPSPSTQQFMMPRSSSRC SCCDGKEPQTITTQLTHHQSKRIKFFEEFE QEKKYRPSHGDKTSNPSVLKWMDIAAGGK QLKELKLKSEQGSAPKGPPRILICEQPTVP RENGKPEAAGPEPSSSGEETPDAALTCLKERR EQLPPQEDSKVTKQDKNLIKPLYDRYRIIKQIL STPSLPTTVSQDTCMLLIKCTDV RENGKPEAAGPEPSSSGEETPDAALTCLKERR EQLPPQEDSKVTKQDKNLIKPLYDRYRIIKQIL STPSLPTTVSQDTCMLLIKTLYDYRIIKQIL STPSLPTTVSQDTCMLLIKTLYDYRIIKQIL STPSLPTTVSQDTCMLLIKTLYDYRIIKQIL ENNRSAACKSPGTGDFSRNSNASNKSVDY SRSQCSCGSLSSQYDYSEDFLCDCSEKANRN YLKQPVVKEKEKKYNVSKSGGQKEISV EKHTWNASLFNSQHMAQRRDAMAHRLS ARLHKKGLKSRPGTGDFSRNSNASNKSVDY SRSQCSCGSLSSQYDYSEDFLCDCSEKANRN YLKQPVVKEKEKKKYNVSKSGGQKEISV EKHTWNASLFNSQHMAQORDAMAHRLS ARLHKKGLKSEPTGDFSRNSNASNKSVDY KNLRQLLRKSQEKERTLSKKLREITSQLLKT KDIDAJACKLSEBOKNLAEBELTIKKLSIITTK MDANDKKIQSLEKQLRINCRAFSRQLKKT KDIDAJACKLSEBOKNLAEBELTIKKLSIITTK MDANDKKIQSLEKQLRINCRAFSRQLKKT KULAAQTATKTIQVEVKHLQOKLKEDREL EKNIYSHRILKNILDITEDVPKVSSTKSVQAD RKILPTSMHQFOTQKSDVPHTYRGKKATO NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGERHILEVQILLENTGRQKDKKEDGEK KNIFVKEEQELPFKILEVHIEPESNQEDVLVR EKFKRSMQRNQVDDTLGKGTAPYTKIGKKATO NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGERHILEVQILLENTGRQKDKKEDGEK KNIFVKEEQELPFKILLENTGRGKDKKEDGEK KNIFVKEEQELPFKILLHILERESNQEDVLVT EKFKRSMQRNGVDDTLGKGTAPYTKIGKKATO NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGERHILEVQILLENTGRQKDKKEDGEK KNIFVKEEQELPFKILLENTIGGRAPANGMR VSHSTGKHLSNREEMELEHSDSQYEPSGKS SIERKVADTTERDKKSSLMEEPTGSQYTLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSKSPSSTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKIRGRILLQCLIDDVKTLHG NAPAFGTBAASGWOPPTYTISGRAFSARYPPP SRRGYSSHHGPSWKKYSLVGKODGE REFKRSMORNGVBORDLERGVQLS QGGNVVIKVKPPSKSGSAASGALVTCRTNKF RKNNYKWAASSKSPRVARRALSPRTQVALG RKLGSHSVASCAPQLLDGRRVDAGHTIDQPP SGSGGPARPASGPGGARSALVTCRTNKF RKNNYKWVAASSKSSPRVARRALSPPRVAAEN VCKASAGMAMKVERVGLLDDFRVAAEN VCKASAGMAMKVERVGLLDDFRVAAEN VCKASAGMAMKVERVGLLDDFRVAAEN VCKASAGMAMKVERVGLLDDFRVAAEN VCKASAGMAMKVERVGLLDDFRVAAEN	ł	ł	l]	SMEDUUVIA TAI ADSA DESACRES DELLA
GTTAHESPONNAFK.OG.TIVRLIOPRIDORTAT SPKDAPETR.ODI.NEEGA.OVJGVKDRAPAS TOSVLADGTDSADPSPVHKJGGNPEADSAPS DLHSVGTSRLILLYHITDGDNPTAVHRIGSJF SGQSQRFNLDPESAPSPPSTOQPMMPRSSSRC SCCDGREPQTITQLTKHIQSLKRKIRKFEEKFE QEKKYRPSHGDKTSNPSVLKMDLAKGRK QLKELKLI.SEEQGSAPKGPPRILLCEDPTVP RENGKPEAAGPPSSSGGETPDAALTCLKEPR EQLPPQEDSKVTKQDKNLIKPLYDRYRIIKQIL STPSLIPTIVSQDTCMLLLCTDV 840 2190 A 6873 2 2054 FFRFYFSFIRLFAMSLADLTKTNIDEHFFGVAL ENNRRSAACKRSPGTGDFSRNSNASNKSVDV SRSQCSCGSLSSGVDVSEDFLOESEKANRN YLKQPVVKEKEKKXVNVSKISQSKGGKEISV EKKHTWNASLFNSQHMAQRRDAMAHRILS ARLHKIGLIKNELADMHAKLEALITENQFLK QUQLRIILKAIGKYENSQNNLPQIMAKHQNEV KNIRQLLRKSQGKERTLSRKTENGQLAETT KDILQALQKLSEDKNLAEREELTHKLSHITTK MDADKKIGDSKGKERTLSRKTSRQLAETT KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIYSHILKNILDHEDVTPKVSSTKSVQAD RKILPFTSMHQGTOKSSDVPLTITKGKKATO NIDHKEKSTEDHEPHCVNKLPKQEDSKRYV EDLSGEEKHLEVQUILENTGRKQKKKEDQEK KNIFVKEQELPPKIUEVHIPERESNQEDVLVR EKFKSMQRNGVDDTUJGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLASGGPANAGNMR YSHSTGKHLSNREEMELEHSUDSOYEPSFGKS SRIKVKDTTERDKKSSLMEELFGGSGVVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSKISPSPTEGKELFGGSGVVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSKISPSPTEGKAGRAGG RCHSTRAGGGPPVPQQHVLERQVQULS QGGNVVIKKYPFSKSGSASSGKGMSLEEFF DTPWSDQRPREGGEEPPRGQLQPSRPTRARG TCSVEDPLLVQKEPGKPRMVKSVGSVGDSP REPRITYSESVIAVKASFPSSLAVTCRTNKF RKNNYKWVAASSKSSRVARRALSPRVAAEN VCKASAGMAMKVEKPQLLADPEPKRRPATS SKPGSAPSKYWKASSPSSSSSFRWQSEAG SKDHASQLSFVLSRSPGORRALAHSGLIFTIS GETELSAYKVKTRTKURRRGSSLPLOPKKSSL	1					į	A CIEN Y ECCHENCE COLLEGIA (COOLECOTO)
SPKDAFETRQDLNEEEAAQVHGVKZDAPASS TQSVLAUGTDSADPSPYHKJGQNEADSAPE DLHSVGTSRLLLVHITDGDNPTAVRHGCSLF SQQSQRFNLDPSAPSPSTQQFMMPRSSSRC SCCDGKEPQTITQLTKHIQSK.KRKFEEKFE QEKKYRPSHGDKTSNPSVLKWMDDLAKGRK QLKELKLK.SEQGSAPKGPPRSLICEQPTVP RENGKPEAAGPEPSSSGEETDAALTCLERR EQLPPQEDSKVTKGDKNLIKHLYDRYDRYRIKQIL STPSLIPTIVSQDTCMLLLCTDV 840 2190 A 6873 2 2054 FFFFYFSFILPAMSLADLTKTNIDEHFFGVAL ENNRRSAACKRSPGTGDFSRNSNASNKSVDY SRSQCSGCSLSSQYDYSEDFLCDCSEKANRN YLKQPVVKEKEKKXYNVSKSGGQKEISV YLKQPVVKEKEKKKYNVSKSGGQKEISV YLKQPVVKEKEKKKYNVSKSGGQKEISV YLKQPVVKEKEKKKYNVSKSGGQKEISV YLKQPVVKEKEKKKYNVSKSGGQKEISV YLKQPVVKEKEKKKYNVSKSGGQKEISV YLKQPVVKEKEKKKYNVSKSGGQKEISV YLKQPVVKEKEKKKYNVSKSGGQKEISV YLKQPVVKEKEKKKYNVSKSGGQKEISV YLKQPVVKEKEKKKYNVSKSGGQKEISV YLKQPVKEKEKKKYNVSKSGGQKEISV YLKQPVKEKEKKKYNVSKSGGQKEISV YLKQPVVKEKELALDMHHKLALLTENOFLK QLQLRILKAGCKERTLSKKLREITSQLLKT KDILQALQKLSEEKKLANVSKSGGQKEISV KNLRQLLRKSGEKERTLSKKLREITSQLLKT KDILQALQKLSEEKKLANVSKSGGQKEISV KNLRQLLRKSGEKERTLSKKLREITSQLLKT KDILQALQKLSEEKKLANVSKSGGQKEISV KNLRQLLRKSGEKERTLSKKLREITSQLLKT KDILQALQKLSEEKKLANVSKSCGQKEIST KKLLAQTATKTLQVEVKHLQQKLKEEKDREL EKNIYSHRILKNILDITEDYPKYSSTSKSVQAD RKLIPFTSMRHQGTQKSSDVPPLTTKGKKATA NDHKEKSTEDHEIPHICVNKLPKOPELOTKRKY EDLSGEKKHLEVQILLENTGRQKDKKEDGEK KNIFVKEEQELPFKILEVHIFRESNQEDVLVR EKKRSMQRNGVDDTLGKGTAPYTKGKATA NDHKEKSTEDHEIPHICVNKLPKOPELOTK RKHYSTETATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHSDSGYPEPSGKS SRIKKVDTTERDKKSSLMEELEJGGGYVLKTD QSSPOVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSKKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAGTTAAASGWOPPTYTISGRAFSARYPP SRRGYSSHHGPSWKKYSLVGXODSP REPRITAGRAFSARYPP SRRGYSSHHGPSWKKYSLVGXODSP REPRITAGRAFSARYPP SRRGYSSHHGPSWKKYSLVGXODSP REPRITAGRAFGARAERSLVYTCRTNKF RKNNYKWVAASSKSSPVARRALSPRVAAEN VCKASAGMANKVEKPOLLDDFYRLATS SKPGSAPSKYKWKASSPSSSSSFRWQSEAG SKDHASQLSFVLSRSPSGDRPALAHSGLIKPLS GETLISAYKVKKTIKURRRGRGSLLEPISDKKSG GETLISAYKVKKTIKURRRGRGSLEPIGLKSG GETLISAYKVKKTIKURRRGRGSLEFIGK	i					İ	
TQSVLAUGTDSADPSPVHKDGQNEADSAPE DLHSVGTSRILL/WHITDGDNPTAVPHGGSLF SGQSQRFNLDPESAPSPPSTQGFMMPRSSSRC SCGDGKEPQTITQLTKHIQSLKKKIKFEEKFE QEKKYRRSHDGKTSNPEVLKWMMDLAKGRK QLKELKLKLSEEQGSAPKGPFRNLICEQPTVP RENGKPEAAGPPESSGEBEFDAALTCLKERR EQLPPQEDSKVTKQDKNLIKPLYDRYRIKQIL STPSLIPTIVSQDTCMLLKCTDV RENGKPEAAGPEPSSGEBEFDAALTCLKERR EQLPPQEDSKVTKQDKNLIKPLYDRYRIKQIL STPSLIPTIVSQDTCMLLKCTDV RENGKPEAACKRSPGTGDFSRNSNASNSVDY SRSQCSCGSLSSQVDYSEICLCOSEKANNR YLKQPVVKEKEKKKYNNSKISQSKGQKEISV EKKHTWNASLFNSQHMARDAMARDLS ARLHKIKGLKNELADMHHKLEAILTENQFLK QLQLRHLKAIGKYENSQNNLPQIMAKHQNEV KNLRQLLRKSQEKERTLSRKLRETDSQLLKT KDILQALQKLSEDKNLABREELTHKLSIITTK MDADKKKQSLEKQLELNGERAFSRQLABTR KTLAAQTATKTLQVEVKHLQQKLKEEDREL EIKNYSHRILKNLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPLTITKGKKATO NIDHKEKSTENHEIPHCVNKLPRQEDSKRXY EDLSGEEKHLEVQILLENGKQKKKEEDDSLVLVR EKKRSMQRNGVDDTLAGKGDKKEEDQEV KNIFVKEEQELPPKIIEVHPERESNQEDVLVR EKKRSMQRNGVDDTLAGKGDKKEEDQEV KNIFVKEEQELPPKIIEVHPERESNQEDVLVR EKKRSMQRNGVDDTLAGKTAPTTKCPLQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLEVGILLENGKGKNEEDDVLVR EKKRSMQRNGVDDTLAGKTAPTTKCPLQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLEVGILLENGKGKNEEDDVLVR EKKRSMQRNGVDDTLAGKATAPTTKCPLQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLEVGILLENGKSKIII 841 2191 A 6874 3 2867 SSRTEMEEKELIRQRILLQGLIDDYKTLHG NAPAPGTPAASGWQAFSKIII 842 2191 A 6874 3 SSRTEMEEKELIRQRILLQGLIDDYKTLHG NAPAPGTPAASGWQAFSKIII 843 2191 A 6874 3 SSRTEMEEKELIRQRILLQGLIDDYKTLHG NAPAPGTPAASGWQAFSKIII 844 2191 A 6874 3 SSRTEMEEKELIRQRILLQGLIDDYKTLHG NAPAPGTPAASGWQAFSKIII 845 SSRTEMEEKELIRQRILLQGLIDDYKTLHG NAPAPGTPAASGWQAFSKIII 846 SSRTEMEEKELIRQRILLQGLIDDYKTLHG NAPAPGTPAASGWQAFSKIII 847 SSRTEMEEKELIRQRILLQGLIDDYKTLHG NAPAPGTPAASGWQAFSASGAAGAGLEEFE DTFWSDQRTREGEGEPPRGQLQPSRTTARG TCSVEDPLLVCQKEPGFRGMVSSVGSVQDSP REPRRTVSESVIAVASSPSSASSSSFRWQAEA VCKASAGMANKVEKPQLUPEPKPRRPATS SKPGSAPSKYXWAASSPSASSSSFRWQSEAG SKDHAPASGFRQARSASLVVTCRTNKF RKNNYKWVAASSSSSSFRWARALSPRVAAEN VCKASAGMANKVEKPQLUPEPKPRRPATS SKPGSAPSKYXWKASSPSASSSSSFRWQSEAG SKDHARGULFFRENGESTSLPGDKKSG	1		1				
DLHSVGTSRLILL/HITDGONPTAVHGCSL/F SGGQSQREN_DPESASPSTQCPMMPRSSSRC SCGDGKEPQITTQLTKHIQSLKRKIRKFEEKFE QEKKYRPSHODKTSNPEVLKWMMDLAKGRK QLKELKLKJSEQGSAPFONLLCEQPTVP RENGKPEAAGPEPSSGEETPDAALTCLKERR EQLPPQEDSKVTKQDKNLIRLYDTRYRIIKQIL STPSLIPTIVSQDTCMLILLCTDV 840 2190 A 6873 2 2054 FFRFYTSFIRLFAMSLADLTKTNIDEHFFGVAL ENNRSAACKRSPGTOBESRNSNASNKSVDY SRSQCSCGSLSSQVDYSEDFLCDCSEKANRN YLKGPVVKEKEKKYNVSKISQSKGQKEISV EKKHTWNASLFNSQIHMAQRRDAMAHRILS ARLHKIKGLKNELADMHHKLEALTERQFLK QLQLRHLKAIGKYENSQNLYNGDMAKQNEV KNLRQLLRKSQEKERTLSRKLRETDSQLLKT KDIQAALQKLSEDKNLAERLHTMLSDUTLYR KDIQAALQKLSEDKNLAERLHTMLSDUTLYR KDIQAALQKLSEDKNLAERLHTMLSDUTLYR KDIQAALQKLSEDKNLAERLHTMLSDUTLYR KDIQAALQKLSEDKNLAERLHTMLSDUTLYR KRILAAQTATKTLQVEVKHQQLCIKEKDREL EIKNYSHRILKNLHITEDYPKVSSTIKSVQAD RKLPFTSMRHQGTQKSDVPPLTTKGKKATO NIDHKEKSTEINHEIPHCVHLKYGEDSKRY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPPKIIEVHPERESNQEDVLVR EKKRSMQRRQVDDTULGKGTAPYTKCPLRQ RRHYSFTEATENLHHGILPASGGRANAGNMR YSHSTIGKHLSNREEMELENSGGYPVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQATSHA FGDSKVTVVNSIKPSSPTEGKRIII 841 2191 A 6874 3 2867 SSRTERMEEKEURRQIRLLQGLIDDVKTLHG NAPAGTFAASGWQPTVQHVLERQVQLS QGQNVVIKVKPPSKSGSAASGAQRGSLEEFE DTPWSDQRPREGEGEPPRQQLPSPTRARG TCSVEDPILLVCQKEPGKPRVSKYSGVSQDSP REPRRTVSESVIAVKASFPSSAJEPRTGVALG RKLOSHSVASCAPQLLGDRRVDRVSGVSQDSP REPRRTVSESVIAVKASFPSSAJEPRTGVALG RKLOSHSVASCAPQLLGDRRVDRAGLEFF DTPWSDQRPREGEGEPPRQQLPSPTRARG TCSVEDPILLVCQKEPGKPNKSVSGVSQDSP REPRRTVSESVIAVKASFPSSAJEPRTGVALG RKLOSHSVASCAPQLLGDRRVDRAGLEFF DTPWSDQRPREGEGEPPRQQLPSPTRARG TCSVEDPILLVCQKEPGKPNKSVSGVSQDSP REPRRTVSESVIAVKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSDARSSSSSFRWQSEAG SKDHASQLSPVLSRSPSDARSSSSSFRWQSEAG SKDHASQLSPVLSRSPSDARSSSSSSFRWGEAG SKDHASQLSPVLSRSPSDDRPALAHSIGKPLS GETPLSAVKVKTRTKLIRRRGSTSLPGDKKSG GETPLSAVKVKTRTKLIRRRGSTSLPGDKKSG	Í I						
SGQSQRFNLDPESAPSPFTQQFMWRSSRC SCGDGKEPQTITQLTKHQSLRKKIRKFEKFE QEKYRPSHGDKTSNPEVLKWMNDLAKGRK QLKELKLKLSEEQGSAPKGPPRNLLCEQPTVP RENGKPEAAGPEPSSSGEBETPDAALTCLKERR EQLPPQEDSKVTKQDKNLIKELYDYRRIKQLL STPSLIPTIVSQDTCMLLLCTDV 840 2190 A 6873 2 2054 FFRFYSFIRLFAMSLADLTKTNIDEHFFGVAL ENNRSAACKRSPGTGDFSRNSNASNKSVDY SRSQCSCGSLSSQVDYSEDPLCDCSEKANRN YLKQPVVKEKEKKKYNVSKISQSKGQKEISV EKKHTWNASLPNSQHMMAQRDAMAHRILS ARLHKIKGLKNELADHHNLLEALTENQFLK QLQLRHLKARGFYENSQNHLPQMAKHQNEV KNLRQLLRKSQEKERTLSRKLRETDSQLKK KDILQALQKLSEDKNLABREELTHKLSHITTK MDANDKKJQSLEKQURLNPJMAKHQNEV KNLRQLLRKSQEKERTLSRKLRETDSQLKK KTLAAQTATKTLQVEVKHLQQKLKEDREL EIKNIYSHRIKNLHDTEDYFKVSSTKSVQAD RKLPFTSMRHQGTQKSDVPPLTTKGKKATO NIDHKEKSTEINHEPPHCVNKLPKQEDSKKY EDLSGEEKHLEVQILLENTIGRQKDKEDQEK KNIFVKEEQELPFKIEVHPERESNQEDVLVR EKFKRSMQRGVDDTLGGTAPTTKGPLRQ RRHYSFTEATENLHHGIPASGGPANAGNMR YSHSTGKHLSNREEMELEHSDSQYEPSFGKS SRIKVKDTTFRDKKSSLMEELPGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPSQASTSHA FGDSKVTVVNSIKPSSTFERKKIII 841 2191 A 6874 3 2867 SSRTEMEEKEILRQIRLLQGLIDDYKTHG NAPAGGTAAGGOPPVPQQHVLERQVQLS QGGNVJKKVKPPSKSGSAASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLOPSTRARG TCSVEDPILLVOCKEPGERPRGLOPSTRARG TCSVEDPILVOCKEPGERPRGAFSASTVAGE QGNVVKKVEPSKSGSAASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLOPSTRARG TCSVEDPILVOCKEPGERPRAGVSVGODSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSSVGGPARASGSPVARRASLSPRVAAEN VCKASAGMANKVEKPQLUEPEPKPRATS SKPGSAPSKYWWAASSSSSFRVARALSSPRVAAEN VCKASAGMANKVEKPQLUEPEPKPRAPATS SKPGSAPSKYWWAASSSSSFRVARRALSSPRVAAEN VCKASAGMANKVEKPQLUEPEPKPRAPATS SKPGSAPSKYWWAASSSSSFRVARALSSPRVAAEN VCKASAGMANKVEKPQLUEPEPKPRAPATS SKPGSAPSKYWKASSPSASSSSSFRWQSEAG SKDHASQLSFVLSRSPSGORRPALAHSIGKPLS GETPLSAVKVKTRTKLERRRGTSTSLPGDKKSG							1QSVLAUGIDSADPSPVHKDGQNEADSAPE
SCGGGKEPQITTO_UTKHIQSLKRKIRKFEKKFE QEKKYRFAHGDKTSNPSUKWAMDID_AKGRK QLKELKLKLSEEQGSAPKGPPRNLLCEQPTVP RENGKYEAAGPEPSSSGEITPDAALTCLKERR EQLPPQEDSKVTKQDKNLIKELYDYRYRIKQIL STPSLIPTIVSQDTCMLLLCTDV FFRFYFSFRLFAMSLADLIKTNIDEHEFGVAL ENNRSAACKRSPGTGDFSRNSNASNKSVDY SRSQCSCGSLSSQVDYSEDFLCDCSEKAINRN YLKQPVVKEKKKKYNYSGSKGQKEISV EKKHTWNASLFNSQHMAQRRDANAHRILS ARHKIKGLKNELADMHKLEAILTENOFLK QLQLRHLKAIGKYENSQNNLPQIMAKRGNEV KNLRQLLRKSQEKERTLSRKLREITDSQLLKT KDIQALQKLSEDKNLAEELTHKLSIITTK MDANDKKIQSLEKQLEILNGRAFSRQLAETR KTLAQTTATKTQVEVEKLPKQEDSKRYY EDLSGEKHLEVQILLENTGRQKDKEEDQEK KNIFVKEEGLEPKILEVIHLKPQEDSKRYY EDLSGEKHLEVQILLENTGRQKDKEEDQEK KNIFVKEEGLEPKILEVIHLKPQEDSKRYY EDLSGEKHLEVQILLENTGRQKDKEEDQEK KNIFVKEEGLEPKILEVIHLPPGNAGTHANGNAM YSHSTIGKHLSNREEMELEHSUSGVYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEPLQSKSGEPLGSGSVEPLTYBGGRATSHAP FGDSKVTVVNSIKPSSPTEGKRKII 841 2191 A 6874 3 2867 SSRTREMEEKELLRRQIRLLQGLIDDYKTHG NAPAGTFAASGWPPTYHSGRAFSARYPP SRRGYSSHIGPSWRKYSLVNRPPGSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASAGAQRGSLEEFE DTFWSDQRFREGEGEPTRGQLOPSTRARG TCSVEDPILLVCQKEPGKPMWKSYGSVGDSP REFRRTVSESVIAVKASFPSSALPPRTGVALG RKLOSHSVASCAPQLLGDRRVDAGHTDQFVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLADPEPKPRPATS SKPGSAPSKYRWKASSPSASSSSSSFRWGEAG SKDHASQLSPVLSRSPSGDRPAALAHSGLKPLS GETPLSAVKVKTRTKLIRRRGSTSLPGDKKSG]				·-		DLHSVGTSRLLL/YHTDGDNPTAVRHGCSL/F
QEKKYRPSHGDKTSNPEVLKWMDLAKGRK QLKELKIKLSIEEQGSAPKGPPRILLCEQPTVP RENGKPEAAGPEPSSSGEETPDAALTCLKERR EQLPPQEDSKVTKQDKNLIKPLYDRYRIKQIL STPSLIPTIVSQDTCALLCTDV 840 2190 A 6873 2 2054 FFRPYFSFRILFAMSLADLTKTNIDEHFFGVAL ENNRSAACKRSPGTGDFSRNSNASNKSVDY SRSQCSGSILSSQYDYSEDFLCDCSEKANRN YLKQPVVKEKEKKYYVSKISOSKGQKEISV EKKHTWNASLFNSQHMIAQRRDAMAHRILS ARLHKIKGLKNELADMHHKLEAILTENQFLK QLQLRHLKAIGKYENSQNNLPQIMAKRQNEV KNLRQLLKKSQEKEFTLSKKLRETDSQLKT KDILQALQKLSEDKNLAERELTHKLSIITTK MDANDKKIOSLEKQLINCRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNYSHRILKNLHDTEDYPKVSSTKSVQAD RKILPFTSRNRQGTQKSDVPPLTTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPPKILEVHPERESNQEDVLVR EKFKRSMQRNGVDDTULGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLASGGPANAGNMR YSHSTGKHLSNREMELEHSIDSGYEPSFGKS SRIKVKDTIFRIKKSSLMEELFGGSYVLKTD QSSSFGVAKGSEFFLQSKSSHPLPPSQASTSHA FGDSKVTVVNISKPSSPTEGKKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIELLGGLIDDYKTLHG NAPAGTTAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLNRPPGPSDPPA DHAVPRI-HGARGGGPPNGQLOGYRPTEARG TCSVEDPLLVCQKEPGKPMVKSVGSVODSP REPRRTYSESVLAVKASFPSSALGPRTAARG TCSVEDPLLVCQKEPGKPMVKSVGSVODSP REPRRTYSESVLAVKASFPSSALGPRTAARG TCSVEDPLLVCQKEPGKPMVKSVGSVODSP REPRRTYSESVLAVKASFPSSALSPRTAARG TCSVEDPLLVCQKEPGKPMVKSVGSVODSP REPRRTYSESVLAVKASFPSSALSPRTAARG TCSVEDPLLVCQKEPGLADPREVALSPRVAARE VCKASAGGMANVEKEPGLADPREVALSPRVAARE VCKASAGGMANVEKEPGLADPREVARAEN VCKASAGGMANVEKEPGLADPREVARAEN VCKASAGGMANVEKEPGLADPREVARAEN VCKASAGGMANVEKEPLADPREVARAEN VCKASAGGMANVEKEPGLADPREVARAEN VCKASAGGMANVEKEPGLADPREVARAEN VCKASAGGMANVEKEPGLADPREVARAEN VCKASAGGMANVEKEPGLADPREVARAEN VCKASAGGMANVEKEPGLADPREVARAEN VCKASAGGMANVEKEPGLADPREVARAEN VCKASAGGMANVEKEPGLADPREVARAEN VCKASAGGMANVEKEPGLADPREVARAEN VCKASAGGMANVEKEPGLADPREVARAEN VCKASAGGMANVEKEPGLADPREVARAEN VCKASAGGMANVEKERGSTSLFOOKKSG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAVRVKTRIKURGSTSLFOOKKSG						į	SGQSQRFNLDPESAPSPPSTQQFMMPRSSSRC
QLKELKILKI.SEEQGSAPKGPPRNILCEQPTVP RENGKPEAAGPEPSSSGEETPDAALTCLKERR EQLPPQEDSKVTKQDKNLIKPI.YDRYNIKQIL STPSLPTIVSQDTCMLLLCTDV							SCGDGKEPQTITQLTKHIQSLKRKIRKFEEKFE
RENGKPEAAGPPSSSGETPDAALTCLKERR EQLPPQEDSKVTKQDKNLIKPLYDRYRIKQIL STPSLIPTIVSQDTCMLLCTDV 840 2190 A 6873 2 2054 FFRFYFSFRIFAMSLADLTKTNIDEHFGVAL ENNRSAACKRSPGTSSRNSNASNKSVDY SRSQCSGSLSSQYDYSEDFLCDCSEKAINRN YLKQPVVKEKKKYVVSKISQSKGQKEISV EKKHTWNASLFNSQHMIAQRRDAMAHRILS ARLHKIKGLKNELADMHHKLEALITENQFLK QLQLRHAAGKYEBCHMIAQRRDAMAHRILS ARLHKIKGLKNELADMHHKLEALITENQFLK COLQLRHAAGKYEBCKHALTENQFLK COLQLRHAAGKYEBCKHALTENQFLK KDILQALQKLSEBCKNLAEREELTHKLSIITTK MDANDKKIQSLEKQLRINCRAFSRQLAETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNYSHRILKNLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGESKHLEVQILLENGSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGESKHLEVQILLENGSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY ENSTGEKHLEVQILLAGRGGPANAGNMR YSHSTGKHLSNREEMELEHSDSQVEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKII 841 2191 A 6874 3 2867 SSRTREMEEKELRGURLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLNMRPGFSDPPA DHAVRPLHGARGGGPPVQQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REFRRTVSESVLAVKASFPSSALGPROVALS GROWVIKVKPPSKSGSASSASGAQRGSLEFFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKRPMVKSVGSVGDSP REFRRTYSESVLAVKASFPSSALPRTGVALG RKLGSHSVASCAPQLLQDRRVDAGHTDOPPY SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPVAARALSPRVAAEN VCKASAGMANKVEKJADPEFKPRATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDNPALAHSGLKPLS GETPLSAYRVKTRTKURRGSTSLFODKKSG	1						
EQLPPQEDSKVTKQDKNLIKPLYDRYRIKQIL STPSLIPTIVSQDTCMLLLCTDV 840 2190 A 6873 2 2054 FFRFYFSFIRLFAMSLADLTKTNIDEHFFGVAL ENNRRSAACKRSPGTGDFSRNSNASNKSVDY SRSQCSCGSLSSQVSEDFLCDCSEKAINRN YLKQPVVKEKEKKYNVSKISQSKGQKEISV EKKHTVWASLFNSQIHMIAQRRDAMAHRILS ARLHKIKGLKNELADMHHKLEALLTENQFLK QLQLRHLKAJGKYENSQNNLPQIMAKHQNEV KNLRQLLRKSQEKERTLSRKIRETDSQLLKT KDILQALQKLSEBKOLAEREELTHKLSIITTK MDANDKKIQSLEKQLRLNCRAFSRQLAETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIYSHRILKNLLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPL/TTKGKKATO NDHKEKSTEINHEIPCVMKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPFKLIEVHIFPERSNQEDVLVR EKFKRSMQRNGVDDTLGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELHSDGSYEPSFOKS SRIKVKDTTFRDKRSSLMEELFGSGYVLKTD QSSPGVAKGSEPLQSKESHPLPSQASTSHA FGDSKVTVVNSIKPSSTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKELIRRQIRLLQGLIDDYKTLHG NAPAFGTFAASGWQPPTYISGRAFSARYPR SRGYSSHHGPSWKKYSLVNRPPGPSDPPA DHAVRPLHGARGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEFFE DTPWSDQRFREGGEPPRGQLQPSPTFARG TCSVEDPLLVCQKEPGKPMVKSVGSVGDSP REFRRTVSESVIAVKASFFSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWAASSKSPRASASLSPRTGVALG SKDHASQLSPVLSRSPSAGSRSSFFRWQSEAG SKDHASQLSPVLSRSPSAGSRSSFFRWQSEAG SKDHASQLSPVLSRSPSAGSRSSFFRWQSEAG SKDHASQLSPVLSRSPSAGSRSLIPRISPLGKKSG				1			
840 2190 A 6873 2 2054 FFRFYFSIRLFAMSLADLTKTNIDEHFFGVAL ENNRRSAACKRSPGTGDFSRNSNASNKSVDYY SRSQCSCGSLSSQVDYSEDFLCDCSEKANRN YLKQPVVKEKKKKYNVSKISQSKGQKEISV EKKHTWNASLFNSQHMIAQRRDAMAHRILS ARLHKIKGLKNELADMHIKLEAILTENQFLK QLQLRILKAIGKYENSQNNLPQIMAKHQNEV KNLRQLLRKSQEKERTLSRKLRETDSQLLKT KDILQALQKLSEDKNLAEREELTHKLSIITTK MDANDKKIQSLEKQLINCRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIVSHRILKNI.HDTEDYPKVSSTKSVQAD RRILFFTSMRHJGCTRSDVPPLTTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEGQELPKUHFBRERSDGDVLVR EKFKRSMQRNGVDDTLGKGTAPYTKGPLRQ RRHYSFTEATENLHHIGLPASGGPANAGNMR YSHSTGKHLSNIEGHELSHDSOYPESFGKS SRIKVKDTTFPDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKELERQGILLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHISGRAFSAYPPP SRRGYSSHIGPSWRKYSLVNRPPGFSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASAGAQRGSLEEFE DTPWSDQRRFEGGEPPRGQLQPSEPTFARG TCSVEDPLLVCQKEPGFRFMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGRQARFASLVVTCRTNKF RKNNYKWAASSKSPRVARRALSPRVAAEN VCKASSAGMANKVEKPQLLADPEFKPRKPATS SKPGSAPSKVKWKASSSPSSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS SKPGSAPSKVKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG	ł						RENGKPEAAGPEPSSSGEETPDAALTCLKERR
840 2190 A 6873 2 2054 FFRFYFSFIRLFAMSILADLTKTNIDEHFFGVAL ENNRRSAACKRSPGTGDFSRNSNASNKSVDY SRSQCSCOSLSSQVDYSEDFLCDCSEKAINRN YIKQPVVKEKEKKYNVSKIQSKGQKEISV EKKHTWASI-FNSQHMIAQRRDAMAHRILS ARLHKIKGLKNELADMHHKLEALITENQFLK QLQLRHLKAIGKYENSQNINQIMAKHQINEV KNLRQLLRKSQEKERTLSRKLRETDSQLLKT KDILQALQKLSEDKNLAEREELTHKLSIITTK MDANDKKIQSLEKQLRLNCAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EEKNIYSHRILKNLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPL/TTKGKKATG NIDHKEKSTENHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDDEK KNIFVKEEQELPPKIIEVHPERESNQEDVLVR EKFRSMQRNGVDDTLGKGTAPYTKGPLRQ RRHYSFTEATENLHIGLPASGGPANAGNMR YSHSTGKHLSNREELEHSDSGYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAGGTPAASGWQPPTVISGRAFSARYPRP SRGYSSHIHGPSWRKYSLVNRPPGSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKYPSKSGSASASAQAGSLEFFE DTPWSDQRFEGGEPPRGOLQPSPTTARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTOVALG RKLOSHSVASCAPQLIGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPANSASSSSFRWQSEAG SKDHASQLSPVLSRSPSAGLAPKTOVALG RKLOSHSVASCAPQLIGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSAGLKPLSSALFV							
ENNRSAACKRSPGTGDFSRNSNASNKSVDY SRSQCSCGSLSSQYDYSEGLAINRN YLKQPVVKEKKKYNVSKISQSKGQKEISV EKKHTWNASLFNSQIHMIAQRRDAMAHRILS ARLHKIKGLKNELADMHHKLEALITENQFLK QLQLRHLKAIGKYENSQNNLPQDMAKHQNEV KNLRQLLRKSQEKERTLSRKLRETDSQLLKT KDILQALQKLSEDKNLAEREELTHKLSIITTK MDANDKKIQSLKQLRINCRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIYSHRILKNLHDTEDYFKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EELSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEGELPPKIIEVHPERESNQEDVLVR EKFKRSMQRNGVDDTUGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHSUSGYFESFGKS SRIKVKDTTFRDKKSSLMEELFGGGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTEMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPR SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRJ-HGARGGQPPVQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEFFE DTPWSDQRPREGGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASACAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSSRWQSEAG SKDHASQLSPVLSRSPGGDRPALAHSGLKPLS SKPGSAPSKYKWKASSPSASSSSSRRWQSEAG SKDHASQLSPVLSRSPGGDRPALAHSGLKPLS GETPLSAYKVKTRKUIRRRGSTSLPGDKKSG							STPSLIPTIVSQDTCMLLLCTDV
ENNRRSAACKRSPGTGDESRNSNASNKSVDY SRSQCSCGSLSSQYDYSEDFLCDCSEKANRN YLKQPVVKEKEKKYNVSKISQSKGQKEISV EKKHTWNASLPNSQHIMIAQRRDAMAHRILS ARLHKIKGLKNELADMHKILEAILTENQFLK QLQLRHLKAIGKYENSQNNLPQIMAKHQNEV KNLRQLLRKSQEKERTLSRKLRETDSQLLKT KDILQALQKLSEDKNLAEREELTHKLSHTIK MDANDKKIQSLEKQLRLNCRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNYSHRILKNLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTENHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQHLENTGRQKDKKEDQEK KNIFVKEEQELPPKIJEVHPERESNQEDVLVR EKFKRSMQRNGVDDTULGKGTAPYTKGPLRQ RRHYSFTEATENLHIGLPASGPANAGNMR YSHSTGKHLSNNEEMELEHSUSOYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEFLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAGTPAASGPVQPTYHSGRAFSARYPR SRRGYSSHIGPSWKKKYSLVNRPPGPSDPPA DHAVRPLHGARGQPPVQHVLERQVQLS QGQNVVKVRPPSKSGSASASGAQRGSLEFE DTPWSDQRPREGGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPQAREASLVVTCRTNKF RKNNYKWVASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPGDDRPALAHSGLKPLS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPGDDRPALAHSGLKPLS GETPLSAYKVKTKTKIKIRRRGSTSLPGDKKSG	840	2190	A	6873	2	2054	FFRFYFSFIRLFAMSLADLTKTNIDEHFFGVAL
SRSQCSGSLSSQYDYSEDFLCDCSEKANRN YLKQPVVKKEKKKYNVSKISQSKGQKEISV EKKHTWNASLFNSQIHMIAQRRDAMAHRILS ARLHKIKGLKNELADMHHKLEAILTENQFLK QLQLRHLKAIGKYENSQNNLPQIMAKHQNEV KNLRQLLRKSQEKERTLSRKLRETDSQLLKT KDILQALQKLSEDKNLAEREELTHKLSIITTK MDAADKKIQSLEKQLRLNCRAFSRQLAIETR KTLAAQTTAKTTLQVEVKHLQQKLKEKDREL EIKNIYSHRILKNLHDTEDYPKVSSTKSVQAD RRILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR EEFKRSMQRNGDTULGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHSUSGYEPSFGKS SRIKVKDTTFRDKKSSLMEELPGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPFSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHISGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGQCPPVPQQHVLERQVQLS QGQNVVKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMYKSVGSVGOSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVASKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRATS SKPGSAPSKYKWKASSSSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSSSPRSGORPALAHSGLKPLS GETPLSAYKVKTRTIKIRRRGSTSLPGDKKSG	1 1						
YLKOPVKEKEKKYNVSKISOSKGOKEISV EKKHTWNASLFNSQIHMAQRRDAMAHRILS ARLHKIKGIKNELADMHIKLEAILTENOPLK QLQLRHLKAIGKYENSQNNL.PQIMAKHQNEV KNLRQLIRKSQEKERTLSRKLETDSQLLKT KDILQALQKLSEDKNL.EAREELTHKLSIITIK MDANDKKIQSLEKQLRI.NCRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNYSHRILKNLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEDNIEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEQQEK KNIFVKEEQELPPKIIEVHPPRESNQEDVLVR EKFKRSMQRNGVDDTLIGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHSJDSGYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRQIRLLQGLIDDYKTLHG NAPAPGTPASGOPANAGNOPTYTHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVQQULSQQLS QGGNVVIKVRPSKSGSASASGAQRGSLEFEE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGRQARGASLVVTCRTINKF RKNNYKWVASSKSPAVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRATS SKPGSAPSKYKWKASSPSASSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWQEAG SKDHASOLSPVLSRSSPSANSSSSSFRWGEAG SKDHASOLSPVLSRSSPSANSSSSSSSFRWGEAG SKDHASOLSPVLSRSSPSANSSSSSFRWGEAG SKDHASOLSPVLSRSSPSANSSSSSFRWGEAG SKDHASOLSPVLSRSSPSANSSSSSFRWGEAG SKDHASOLSPVLSRSSPSANSSSSSFRWGEAG SKDHASOLSPVLSRSSPSANSSSSSFRWGEAG SKDHASOLSPVLSRSSPSANSSSSSSSSSSSFRWGEAG SKDHASOLSPVLSRSSPSANSSSSSSSSSSSSSSSSFRWGEAG SKDHASOLSPVLSRSSPSANSSSSSSSSSSSSSSSSSFRWGEAG SKDHASOLSPVLSRSSPSANSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS					•		l
EKKHTWNASLFNSQIHMIAQRRDAMÄHRILS ARLHKIKGLKNELADMHIKLEALTENQFLK QLQLRHLKAIGKYENSQNNLPQIMAKHQNEV KNLRQLLRKSQEKERTLSRKLRETDSQLLKT KDILQALQKLSEDKNLAEREELTHKLSIITTK MDANDKKIQSLEKQLRLINCRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIYSHRILKNLHDTDEVYRVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLEATIGRQKDKKEDQEK KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR EKFKRSMQRNGVDDTULGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELLEHSIDSGYPEPSGKS SRIKVKDTITRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKELRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVIKVKYPSKSGSASASGAQRGSLEFFE DTPWSDQRPREGGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCKTNKF RKNNYKWVASSKSPRVARALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQEEAG SKDHASQLSPVLSRSPSGOLRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG							
ARLHKIKGLKNELADMIHIKLEAILTENOFLK QLQLRHLKAIGKYENSQNNLPQIMAKHQNEV KNLRQLLRKSQEKERTLSRKLRETDSQLLKT KDILQALQKLSEDKNLAEREELTHKLSIITTK MDANDKKIQSLEKQLRLNCRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIYSHRILKNLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEERHLEVQILLENTGRQKDKKKEDQEK KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR EKFKRSMQRNGVODDTULGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHSDSOVPPSGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGITAASGWQPPTYTISGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEFFE DTPWSDQRPREGEGEPPRGQLQPSRTTRARG TCSVEDPLLVCQKEPGKPRMYKSVSGVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSRVAAEN VCKASAGMANKVEKPQLIADPEPKPRATS SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGDNRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG	1				•		
QLQLRHILKAIGKYENSQNNLPQIMAKHÖNEV KNLRQLLRKSQEKERTLSRKLRETDSQLLKT KDILQALQKLSEDKNLAEREELTHKLSIITTK MDANDKKIQSLEKQLRLNCRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIVSHRILKNLHDITEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR EKFRSMQRNGVDDTVLGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHSDSGVEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRQRIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWKKYSLVNPRPGFSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPRGQLQPSRFTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKNDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG	1						
KNLRQLLRKSQEKERTLSRKLRETDSQLLKT KDILQALQKLSEDKNLAEREELTHKLSIITTK MDANDKKIQSLEKQLRLNCRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIYSHRILKNLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPPKIIEVIHPERESNOEDVLVR EKFKSSMQRNGVDDTNLGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHSDSGYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSFGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRQIRLLQGLIDDYKTLHG NAPAPGTPASGWQPPTYHSGRAFSARYPR SRRGYSSHHGFSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEFFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVQDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAQQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSFRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG	1 1	1					
KDILQALQKLSEDKNLAEREELTHKLSIITTK MDANDKKIQSLEKQLRINGRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIYSHRILKNLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPPKILEVHPERESNQEDVLVR EKFKRSMQRNGVDDTLGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHSUSGYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAFGTPAASGWQPTYTHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEFFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEFKPRKPATS SKPGSAPSKYKWKASSPSSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG	1		-				WILDUIDEN OUT ENSOUND OF THE OUT OF
MDANDKKIQSLEKQLRLNCRAFSRQLAIETR KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIYSHRILKNLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEGQELPPKIIEVIHPERESNQEDVLVR EKFKRSMQRNGVDDTULGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHSDSGYEPSFGKS SIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVKVKPPSKSGSASSAGAQRGSLEEFE DTPWSDQRPREGGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRWVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRAPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIRRRGSTSLPGDKKSG	1 1						
KTLAAQTATKTLQVEVKHLQQKLKEKDREL EIKNIYSHRILKNI.HDTEDVPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR EKFKRSMQRNGVDDTVLGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHS\DSGYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGQPPVPQQHVLERQVQLS QGQNVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPFKPRATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG							
EIKNIYSHRILKNLHDTEDYPKVSSTKSVQAD RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR EKFKRSMQRNGVDDTTLGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHSDDSOYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEPPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVQQUFVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEFFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRWKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTINKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRRALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG	1 !						WILL A OTATIVT OVERVILL CONTURED IN
RKILPFTSMRHQGTQKSDVPPL/TTKGKKATG NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPPKIIEVHPERESNQEDVLVR EKFKRSMQRNGVDDT\LGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREBMELEHSUDSGYEFSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTTPAASGWOPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPI.HGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMYKSVGSVGDSP REPRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLLADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG							KILAAQIAIKILQVEVKHLQQKLKEKDREL
NIDHKEKSTEINHEIPHCVNKLPKQEDSKRKY EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR EKFKRSMQRNGVDDTLLGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHSUDSGYEFSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPI.HGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEFFE DTPWSDQRPREGEGEPPRGQLQPSRPTARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REFRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLLADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG							
EDLSGEEKHLEVQILLENTGRQKDKKEDQEK KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR EKFKRSMQRNGVDDT\LGKGTAPYTKGPLRQ RRHYSFTEATENLHIGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHS\DSGYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTTAASGWQPPTYHSGRAFSARYPRP SRRGYSSHIGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRQLQPSRTTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWAASSKSFRVARRALSFRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKYGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG		1			j		
KNIFVKEEQELPPKIIEVIHPERESNQEDVLVR EKFKRSMQRNGVDDT\LGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHS\DSGYPESFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKYSLVNRPPGPSDPPA DHAVPLHGARGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG	1 1	- 1	ĺ	1	i		
EKFKRSMQRNGVDDT\LGKGTAPYTKGPLRQ RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHS\DSGYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKYSLVNRPPGPSDPPA DHAVPPLHGARGQQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG		1	1	1			
RRHYSFTEATENLHHGLPASGGPANAGNMR YSHSTGKHLSNREEMELEHSDSGYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSAASGAQRGSLEEFE DTPWSDQRPREGEGPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG		1		-		i	
YSHSTGKHLSNREEMELEHS\DSGYEPSFGKS SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTEMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG			ł	1		İ	
SRIKVKDTTFRDKKSSLMEELFGSGYVLKTD QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVYTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG]				ļ		
QSSPGVAKGSEEPLQSKESHPLPPSQASTSHA FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG		ł	l	ļ	l	ł	
FGDSKVTVVNSIKPSSPTEGKRKIII 841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPI.HGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG				1		l	
841 2191 A 6874 3 2867 SSRTREMEEKEILRRQIRLLQGLIDDYKTLHG NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG				i		l	
NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG	J		.,				
NAPAPGTPAASGWQPPTYHSGRAFSARYPRP SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG	841	2191	A	6874	3	2867	
DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG	[[1	1	1	ļ	J	NAPAPGTPAASGWQPPTYHSGRAFSARYPRP
DHAVRPLHGARGGQPPVPQQHVLERQVQLS QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG		1	1			j	SRRGYSSHHGPSWRKKYSLVNRPPGPSDPPA
QGQNVVIKVKPPSKSGSASASGAQRGSLEEFE DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG]	i	į	ŀ			DHAVRPI.HGARGGQPPVPQQHVLERQVQLS
DTPWSDQRPREGEGEPPRGQLQPSRPTRARG TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG			I	[f		QGQNVVIKVKPPSKSGSASASGAORGSLEFFF
TCSVEDPLLVCQKEPGKPRMVKSVGSVGDSP REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG		ļ		ļ			
REPRRTVSESVIAVKASFPSSALPPRTGVALG RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG		ļ					TCSVEDPLLVCOKEPGKPRMVKSVGSVGDSP
RKLGSHSVASCAPQLLGDRRVDAGHTDQPVP SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG		1	- 1	1	ļ	}	
SGSVGGPARPASGPRQAREASLVVTCRTNKF RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGDRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG	1	l	ļ	- 1			
RKNNYKWVAASSKSPRVARRALSPRVAAEN VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSSFRWQSEAG SKDHASQLSPVLSRSPSGD\RPALAHSGLKPLS GETPLSAYKVKTRTKIIRRGSTSLPGDKKSG		ļ					SGSVGGPARPASGDP CADE AST MUTCHTARE
VCKASAGMANKVEKPQLIADPEPKPRKPATS SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGDVRPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG	ļ .	l	. 1	. !			
SKPGSAPSKYKWKASSPSASSSSFRWQSEAG SKDHASQLSPVLSRSPSGD\RPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG	1 1	ł	l	l	l	ļ	
SKDHASQLSPVLSRSPSGD/RPALAHSGLKPLS GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG		1]		1		•
GETPLSAYKVKTRTKIIRRRGSTSLPGDKKSG					İ	İ	
1 1 1 1 1 1 1 1 1 1			- 1	ŀ	1	j	•
TSPAATAKSHLSLRRRQALRGKSSPVLKKTPN		ļ	l		I		
	L					. <u> </u>	TSPAATAKSHLSLRRRQALRGKSSPVLKKTPN

NO: of nucleotide cotide sequence USSN 09496 (USSN 09496 corresponding to last amino acid residue of peptide sequence USSN 09496 (ISSN 09496 corresponding to last amino acid residue of peptide residue of peptide sequence Sequenc
eotide sequence USSN 09/496 orresponding to last amino acid residue of peptide residue of peptide sequence sequ
sequence 09/496 correspondi 10 1st amino acid residue of peptide criditue of peptide sequence v=1/2 residue of peptide sequence sequen
uence 914 ng to first amino acid residue of peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence peptide sequence periode seque
amino acid residue of peptide sequence peptide sequence sequence
residue of peptide sequence
peptide sequence
Sequence nucleotide insertion
VRTAPTSKVIKTRYBIVKKTPASPLSAPPFPL LPSWRARRLSLSRSLVLNRLRPVASGGGKA PGSPWRSKGYRCIGGVLYKVSANKLSKTS QPSDAGSRPLLRTGRLDPAGSCSRSLASRAV RSLAIIRQARQRREKRKEYCMYYNRFGRCN GERCPYHDPEKVAVCTRPVRGTCKKTDGT PFSHHVSKEKMPVCSYFLKGICSNSNCPYSH YVSRKAEVCSDFLKGYCPLGAKCKKKHTLI PDFARRGACPRGAQCQLLHRTQKRHSRRAA SPAPGFSDATARSRVSASHGPRKPSASQRFT QTPSSAALTAAAVAAPPHCPGGSASPSSSKA SSSSSSSPPASLDHELAPSLQEAALAAACSN LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKD GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDN RNKFKVINVDDDGNELGSGIMELTDTELILY RKRDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFAFKCARAEELFMILQEIMQNNSIN VVEEPVERNNHQTELEVPRTPTTPTGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEBI KNRTSVHYPLEARVSNAESSTFKEEPSSIEDB DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARI LSRDEDDNLGPKTPSLNGYHNNLDPMHNY NTENVTVPASAHKIEYSRRRDCTPTVFNFDI RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTTELYAVIDIERTAAMSNLQKAI PRDDGTSRKTRHNSTUDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLUPPVSRA 447TACVLTTAAVGLELTPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRN
VRTAPTSKVIKTRYBIVKKTPASPLSAPPFPL LPSWRARRLSLSRSLVLNRLRPVASGGGKA PGSPWRSKGYRCIGGVLYKVSANKLSKTS QPSDAGSRPLLRTGRLDPAGSCSRSLASRAV RSLAIIRQARQRREKRKEYCMYYNRFGRCN GERCPYHDPEKVAVCTRPVRGTCKKTDGT PFSHHVSKEKMPVCSYFLKGICSNSNCPYSH YVSRKAEVCSDFLKGYCPLGAKCKKKHTLI PDFARRGACPRGAQCQLLHRTQKRHSRRAA SPAPGFSDATARSRVSASHGPRKPSASQRFT QTPSSAALTAAAVAAPPHCPGGSASPSSSKA SSSSSSSPPASLDHELAPSLQEAALAAACSN LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKD GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDN RNKFKVINVDDDGNELGSGIMELTDTELILY RKRDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFAFKCARAEELFMILQEIMQNNSIN VVEEPVERNNHQTELEVPRTPTTPTGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEBI KNRTSVHYPLEARVSNAESSTFKEEPSSIEDB DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARI LSRDEDDNLGPKTPSLNGYHNNLDPMHNY NTENVTVPASAHKIEYSRRRDCTPTVFNFDI RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTTELYAVIDIERTAAMSNLQKAI PRDDGTSRKTRHNSTUDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLUPPVSRA 447TACVLTTAAVGLELTPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRN
LPSWRARRLSLSRSLVLNRLRPVASGGGKA PGSPWRSKGYRCIGGVLYKVSANKLSKT QPSDAGSRPLLRTGRLDPAGSCSRSLASRAV RSLAIRQARQRREKRKEYCMYYNRFGRCN GERCPYHIDPEKVAVCTRFVRGTCKKTDGT PFSHHVSKEKMPVCSYFLKGICSNSNCPYSH YVSRKAEVCSDFLKGYCPLGAKCKKKHTLL PDFARRGACPRGAQCQLLHRTQKRHSRRAA SPAPGPSDATARSRVSASHGPRKPSASQRPT QTPSSAALTAAAVAAPPHCPGGSASPSSSKA SSSSSSSPASLDHEAPSLQEAALAAACSN LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKD GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNI RNKFKVINVDDDGRELGSGIMBELTDTELLLY RKRDSVKWHYLCLRRYGYDSNLFSESGRR QTGQGIFAFKCARAEELFNMLQEIMQNNSN VVEEPVVERNNHQTELEVPRTPRTPTTPGFA QNLPNGYPRYSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEQVHTYVNTTGVQEEI KNRTSVHVPLEARVSNAESSTPKEEPSSIEDD DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEAR LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYJQVDLEGGSDSDNQTIPKTY TPLPQTPTFRRTELYAVIDLERTAAMSNLQKAI PRDDGTSRKTRHNSTDLPL 843 2193 A 6919 2 663 AGRFGTTHASGKMAYQSLRLEYLQIPPVSK YTTACVLTTAAVQLELITPFQLYFNPELIFKH QUWRLITNFLFFGPVGFNFLFNMIFLYRYCRM
QPSDAGSRPLLRTGRLDPAGSCSRSLASRAV RSLAIIRQARQRREKREYCMYYNRFGRCN GERCPYIHDPEKVAVCTRFVRGTCKKTTDGT PFSHHVSKEKMPVCSYFLKGICSNSNCPYSH YVSRKAEVCSDFLKGYCPLGAKCKKKHTLI PDFARRGACPRGAQCQLLHRTQKRHSRRAF SPAPGPSDATARSRVSASHGPRKPSASQRPT QTPSSAALTAAAVAAPPHCPGGSASPSSKA SSSSSSSPPASLDHEVAPSLQEAALAAACSN LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKD GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNI RNKFKVINVDDDGNELGSGIMELTDTELIL Y RKRDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEBI KNRTSVHVPLEARVSNAESSTPKEEPSSIEDD DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARK LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRADCTFTVFNFDII RPSLEHRQLYNIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRKKTRINSTDLPL 843 2193 A 6919 2 663 AGRPGTTHAAVQSLRLEYLQIPPVSRA YTTACVLITAAVQLELITPFQLYFNPELIEKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRM
RSLAIIRQARQRREKRKEYCMYYNRFGRCN GERCPYTHDPEKVAVCTRFVRGTCKKTDGT PFSHHVSKEKMPVCSYFLKGICSNSNCPYSH YVSRKAEVCSDFLKGYCPLGAKCKKKHTLL PDFARRGACPRGAQCQLLHRTQKRHSRRAA SPAPGPSDATARSRVSASHIGPRKPSASQRPT QTPSSAALTAAAVAAPPHCPGGSASPSSSKA SSSSSSSSPASLDHE\APPHCPGGSASPSSSKA LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKD GKPHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDN RNKFKVINVDDDGNELGSGIMELTDTELILY RKRDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTFTTPGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEEI KNRTSVHPLLAREVSNAESSTPKEEPSSIEDE DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARI LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKA PRDDGTSRKTRINSTDLPL 843 2193 A 6919 2 663 AGRPGTTHAAVQLELITPFQLYGIPPVSRA YTTACVLTTAAVQLELITPFQLYGNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
GERCPYHIDPĖKVAVCTRFVRGTCKKTDGT PFSHHVSKEKMPVCSYFLKGICSNSNCPYSH YVSRKAEVCSDFLKGYCPLGAKCKKKHTILI PDFARRGACPRGAQCQLLHRTQKRHSRRAA SPAPGPSDATARSRVSASHGPRKPSASQRPT QTPSSAALTAAAVAAPPHCPGGSASPSSSKA SSSSSSSSPASLDHEAPSLQEAALAAACSNI LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKD GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEAMGSCCSCPDKDTVPDNI RNKFKVINVDDDGNELGSGIMELTDTELLLY RRXDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTTTTOFFA QNLPNGYPPYSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEEI KNRTSVHVPLEARVSNAESSTPKEEPSSIEDF DPQILLEPEGVKFVLGFTPVQKQLMEKEKLE QLGRDQVSSGANNTEWDTGYDDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARI LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHSGKTRHNSTDLPL 843 2193 A 6919 2 1663 AGRPGTTHSGKTRJCTFTVFNFDLIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRN
PFSHHVSKEKMPVCSYFLKGICSNSNCPYSH YVSRKAEVCSDFLKGYCPLGAKCKKHTLI PDFARRGACPRGAQCQLLHRTQKRKHSRRAA SPAPGPSDATARSRVSASHGPRKPSASQRPT QTPSSAALTAAAVAAPPHCPGGSASPSSSKA SSSSSSSPPASLDHE\APSLQEAALAAACSN LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKD GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDIN RNKFKVINVDDDGNELGSGIMELTDTELILY RKRDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEBI KNRTSVHVPIEARVSNAESSTPKEEPSSIEDB DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARI LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTFKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRKTRHNSTUDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRN
YVSRKAEVCSDFLKGYCPLGAKCKKKHTLI PDFARRGACCQLLHRTQKRHSRRAA SPAPGPSDATARSRVSASHGPRKPSASQRPT QTPSSAALTAAAVAAPPHCPGGSASPSSSKA SSSSSSSSPASLDHEAPSLQEAALAAACSN LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKD GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNI RNKFKVINVDDDGNELGSGIMELTDTELILY RKRDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFARKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPTTPTGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEBI KNRTSVHVPLEARVSNAESSTPKEEPSSIEDB DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARI LSRDEDDDNLGPKTPSLNGYHNNLDPMHNYV NTEENVTVPASAHKIEYSRRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTFKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNSTÜLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRN
PDFARRGACPRGAQCQLLHRTQKRHSRRAA SPAPGPSDATARSRVSASHGPRKPSASQRPT QTPSSAALTAAAVAAPPHCPGGSASPSSKASSSSSSSPPASLDHE\APPLQEAALAAACSN LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKD GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNI RNKFKVINVDDGNELGSGIMELTDTELLLY RKRDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGESTHPLLVAEEQVHTYYNTTGVQEEI KNRTSVHVPLEARVSNAESSTPKEEPSSIEDD DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRGRITSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARL LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRN
SPAPGPSDATARSRVSASHGPRKPSASQRPT QTPSSAALTAAAVAAPPHCPGGSASPSSSKA SSSSSSSPASLDHE\APSLQEAALAAACSNI LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKD GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNI RNKFKVINVDDDGNELGSGIMELTDTELIL Y RKRDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTTGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEEI KNRTSVHVPLEARVSNAESSTPKEEPSSIEDF DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARI LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDIDI RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSNKTRHNSTUDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITFFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRN
QTPSSAALTAAAVAAPPHCPGGSASPSSSKA SSSSSSSPASLDHEVAPSLQEAALAAACSN LCKLPSFISLQSSPSFGAQPRVRAPRAPLTKD GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNI RNKFKVINVDDDGNELGSGIMELTDTELILLY RKRDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEEI KNRTSVHVPLEARVSNAESSTPKEEPSSIEDF DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARF LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRN
SSSSSSSPPASLDHEVAPSLQEAALAAACSNI LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKD GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNI RNKFKVINVDDDGNELGSGIMELTDTELILY RKRDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEEI KNRTSVHVPLEARVSNAESSTPKEEFSSIEDR DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARI LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRN
LCKLPSFISLQSSPSPGAQPRVRAPRAPLTKD GKPLHIKPRL 842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNI RNKFKVINVDDDGNELGSGIMELTDTELILY RKRDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEBI KNRTSVHVPLEARVSNAESSTPKEEPSSIEDB DPQILLEPEGVKFVLGPTPVQKQLMEKEKLL QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARI LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNST\DLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLITTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNI RNKFKVINVDDDGNELGSGIMELTDTELILY RKRDVKWHYLCLRRYGYDSNLFSFESGR QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEBI KNRTSVHVPLEARVSNAESSTPKEEPSSIEDE DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARI LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRKTRHNST\DLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLITAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
842 2192 A 6898 506 2071 WPDLVHTWSSEEAMGSCCSCPDKDTVPDNI RNKFKVINVDDDGNELGSGIMELTDTELILY RKRDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFA QNLPPNGYPRYPSFGDASSHPSSKHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEEI KNRTSVHVPLEARVSNAESSTPKEEPSSIEDE DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARF LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRAYTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
RNKFKVINVDDDGNELGSGIMELTDTELILY RKRDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEBI KNRTSVHVPLEARVSNAESSTPKEEPSSIEDE DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARE LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
RKRDSVKWHYLCLRRYGYDSNLFSFESGRR QTGQGIFAFKCARAEELFNMLQEIMQNNSIN VVEEPVVERNNHQTELEVPRTPRTPTTPGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEEL KNRTSVHVPLEARVSNAESSTPKEEPSSIEDE DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARE LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
VVEEPVVERNNHQTELEVPRTPRTPTTPGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEER KNRTSVHVPLEARVSNAESSTPKEEPSSIEDE DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARE LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNST\DLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
VVEEPVVERNNHQTELEVPRTPRTPTTPGFA QNLPNGYPRYPSFGDASSHPSSRHPSVGSAR PSVGEESTHPLLVAEEQVHTYVNTTGVQEER KNRTSVHVPLEARVSNAESSTPKEEPSSIEDE DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARE LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNST\DLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
PSVGEESTHPLLVAEEQVHTYVNTTGVQEBI KNRTSVHVPLEARVSNAESSTPKEEPSSIEDE DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERROA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARI LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRKTRHNST\DLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLITAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
KNRTSVHVPLEARVSNAESSTPKEEPSSIEDE DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARE LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNST\DLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLITAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
DPQILLEPEGVKFVLGPTPVQKQLMEKEKLE QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARE LSRDEDDNLGPKTPSLNGYHINNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
QLGRDQVSGSGANNTEWDTGYDSDERRDA SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARI LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTTVFNFFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
SVNKLVYENINGLSIPSASGVRRGRLTSTSTS TQNINNSAQRRTALLNYENLPSLPPVWEARI LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNSTDLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
TQNINNSAQRRTALLNYENLPSLPPVWEARI LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSR\KTRHNST\DLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
LSRDEDDNLGPKTPSLNGYHNNLDPMHNYV NTENVTVPASAHKIEYSRRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNST\DLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLITAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
NTENVTVPASAHKIEYSRRRDCTPTVFNFDII RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNST\DLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
RPSLEHRQLNYIQVDLEGGSDSDNPQTPKTP TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSR\KTRHNST\DLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRA YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
TPLPQTPTRRTELYAVIDIERTAAMSNLQKAI PRDDGTSRIKTRHNST\DLPL 843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRAYTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRAYTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
843 2193 A 6919 2 663 AGRPGTTHASGKMAYQSLRLEYLQIPPVSRAYTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
YTTACVLTTAAVQLELITPFQLYFNPELIFKH QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRN
QIWRLITNFLFFGPVGFNFLFNMIFLYRYCRA
LEEGSFRGRTADFVFMFLFGGFLMTLFGLFV
L/VFLGPGLYNN/GSSMCGAE\EPLCPHELLRI
SQLPGPLSALGAHGIFLVVGELNHCGPFGYC
WTHIFFLGRCISQSTWWNKNSENTTYFESYF
844 2194 A 6928 902 366 HRLCMPIQGACGERME/FSLLLPGLECNGVII
AHCNLRLPGSSNSPASASQVAGITGVCHHAR
LIFVFSVETGFLHAGQAGLELLTSGDPPASAS
QSAGITGKSQHTRPGYEFIIPYSAAQEDALKA LM
845 2195 A 6939 1660 317 LYPENLGESLFPILLLPPPWPDGGRPCCVEMS
TRAKKLRRIWRILEEKESVAGAVOTLLLRSO
GGV\TSAAASTLSEPPRRTQESRTRTRALGLP
LPMEKLAASTEPQGPRPVLGRESVQVPDDQI
FRSFRSECEAEVGWNLTYSRAGVSVWVOAV
EMDRTLHKIKCRMECCDVPAETLYDVLHDII
YRKKWDSNVIETFDIARLTVNADVGYYSWR
CPKPLKNRDVITLRSWLPMGADYIIMNYSVK
HPKYPPRKDLVRAVSIQTGYLIQSTGPKSCVI
YLAQVDPKGSLPKWVVNKSSQFLAPKAMKI
MYKACLKYPEWKQKHLVPHFKPWLVHPEQSI
LPSLALS\ELSVQHADS\LENIDESAV\AESREE
RIMGGAGGEGISDDDTSLYAEAPHRFRETETO
PGAGRALGAAAAPALSPLHPPGTWWHRARF
RRVLQPGWTEPQ

Cero in	CEC 15	T 3.7	Loro	7 p. 7	T 5	
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in NO:	beginning nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-		USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline.
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			1	peptide	•	/=possible nucleotide deletion, \=possible
L		ĺ	1	sequence	[nucleotide insertion
846	2196	Α	6944	42	2672	RRKMAGCRGSLCCCCRWCCCCGERETRTPE
			1	1	İ	ELTILGETQEEEDEILPRKDYESLDYDRCINDP
1		ļ				YLEVLETMDNKKGRRYEAVKWMVVFAIGV
1 1		1	Ì	ł		CTGLVGLFVDFFVRLFTQLKFGVVQTSVEECS
i . I				İ		QKGCLALSLLELLGFNLTFVFLESLLGLIEPVE
		l	} .	1	1	AGSGITEGKCYLYARQVPGLVRLPTLLWKAL
]		GVLLTVAAMLLI\GLGSPMIHSGSVVGAGLPQ
		ļ		ł		FQSISLRKIQFNFPYFRSDRYGK\DKRDFVSAG
		ļ		ĺ	ļ	AAAGVAAAFGAPIGGTLFSLEEGSSFWNQGL
)		l	}	ļ	1	TWKVLFCSMSATFTLNFFRSGIQFGSWGSFQL
		1	1			PGLLNFGEFKCSDSDKKCHLWTAMDLGFFV
1		1			j	VMGVIGGLLGATFNCLNKRLAKYRMRNVHP
		ŀ				KPKLVRVLESLLVSLVTTVVVFVASMVLGEC
1		ļ				RQMSSSSQIGNDSFQLQVTEDVNSSIKTFFCP NDTYNDMATLFFNPQESAILQLFHQDGTFSPV
						TLALFFVLYFLLACWTYGISVPSGLFVPSLLC
						GAAFGRLVANVLKSYIGLGHIYSGTFALIGAA
						AFLGGVVRMTISLTVILIEST\NEITYGLPIMVT
]			LMVGKWTGDFFNKGNYDIHVGLRGVPLLEW
						ETEVEMDKLRASDIMEPNLTYVYPHTRIQSLV
						SILRTTVHHAFPVVTENRGNEKEFMKGNQLIS
1			·			NNIKFKKSSILTRAGEQRKRSQSMKSYPSSEL
i l						RNMCDEHIASEEPAEKEDLLQQMLERRYTPY
ĺĺ			[[.		PNLYPDQSPSEDWTMEERFRPLTFHGLILRSQ
1 1						LVTLLVRGVCYSESQSSASQPRLSYAEMAED
						YPRYPDIHDLDLTLLNPRMIVDVTPYMNPSPF
						TVSPNTHVSQVFNLFRTMGLRHLPVVNAVGE
847	2197	A	6951	3	1994	IVGIITRHNLTYEFLQARLRQHYQTI
047	2197	Α.	וכפס	3	1994	NTNSSSVTNSAAGVEDLNIVQVTVPDNEKER
				i	ł	LSSIEKIKQLREQVNDLFSRKFGEAIGVDFPVK
						VPYRKITFNPGCVVIDGMPPGVVFKAPGYLEI SSMRRILEAAEFIKFTVIRPLPGLELSNGEYST
						VGKRKIDQEGRVFQEKWERAYFFVEVONIST
						CLICKRSMSVSKEYNLRRHYQTNHSKHYDQY
						MERMRDEKLHELKKGLRKYLLGLSDTECPE
,						QKQVFANPSPTQKSPVQPVEDLAGNLWEKLR
	ŀ			ļ		EKIRSFVAYSIAIDEITDINNTTQLAIFIRGVDE
				ĺ		NFDVSEELLDTVPMTGTKSGNEIFSRVEKSLK
ļ)			ا ا		NFCINWSKLVSVASTGTPPMVDANNGLVTKL
					}	KSRVATFCKGAELKSICCIIHPESLCAQ\KLKM
]			1	l	DHVMDVVVKSVNWICSRGLNHSEFTTLLYEL
1					ĺ	DSQYGSLLYYTEIKWLSRGLVLKRFFESLEEI
1					ŀ	DSFMSSRGKPLPQLSSIDWIRDLAFLVDMTM
1				-		HLNALNISLQGHSQIVTQMYDLIRAFLAKLCL
1	1		1	1	- 1	WETHLTRNNLAHFPTLKLVSRNESDGLNYIP
	ļ				ŀ	KIAELKTEFQKRLSDFKLYESELTLFSSPFSTKI
1	i	i	1		l	DSVHEELQMEVIDLQCNTVLKTKYDKVGIPE
1				ļ	1	FYKYLWGSYPKYKHHCAKILSMFGSTYICEQ
848			6985	3	289	LFSIMKLSKTKYCSQLKDSQWDSVLHIAT
	2198	A		-	207	SVQYLPGRPTRTHASTDAPLMLKFTPLPSKTK ASAPVQCLLLMAATFSPQGLAKPHSGTIPIT\C
1	2198	A	0,03	ı		ASAF VULLLMAA IPSPUGLAKPHSGTIPITIC
	2198	A			ļ	CENTA DITVIDIODI ECVIDITORIO CONTRATO
849				963	-	CFNAINTKIPIQRLESYTRITNIQCPKEAVM
849	2198	A	6999	963	5	CFNAINTKIPIQRLESYTRITNIQCPKEAVM LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM
849				963	5	CFNAINTKIPIQRLESYTRITNIQCPKEAVM LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP
849				963	5	CFNAINTKIPIQRLESYTRITNIQCPKEAVM LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP
849				963	5	CFNAINTKIPIQRLESYTRITNIQCPKEAVM LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS
849				963	5	CFNAINTKIPIQRLESYTRITNIQCPKEAVM LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT
849				963	5	CFNAINTKIPIQRLESYTRITNIQCPKEAVM LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT TGVLLSLIHLVLLLPLPFCRPQKIYHFFCEILA
849				963	5	CFNAINTKIPIQRLESYTRITNIQCPKEAVM LDFLCHRDMGDNITSITEFLLLGFPVGPRIQM LLFGLFSLFYVFTLLGNGTILGLISLDSRLHAP MYFFLSHL\AVVDIAYACNTVPRMLVNLLHP AKPISFAGRMMQTFLFSTFAVTECLLLVVMS YDLYV\AICHPLRYLAIMTWRVCITLAVTSWT

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	Ì	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
			į	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
			ì	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide		/=possible nucleotide deletion, \=possible
	1	J		sequence		nucleotide insertion
		T				GLFYGTAIIMYVGPRYGNPKEQKKYLLLFHS
			1			LFNPMLNPLICSLRNSEVKNTLKRVLGVERAL
850	2200	A	7001	1	1011	MGNDSVSYEYGDYSDLSDRPVDCLDGACLAI
1	ĺ	1	ĺ		i ·	DPLRVAPLPLYAAIFLVGVPGNAMVAWVAG
		1				KVARRRVGATWLLHLAVADLLCCLSLPILAV
	1				1	PIARGGHWPYGAVGCRALPSIILLTMYASVLL
		l			ĺ	LAALSADLCFLALGPAW\CLRFS/GACGVQVA
	})			J	CGAAWTLALLLTVPSAIYRRLHQEHFPARLO
	l				Ì	CVVDYGGSSSTENAVTAIRFLFGFLGPLVAVA
						SCHSALLCWAARRCRPLGTAIVVGFFVCWAP
	İ		1		·	YHLLGLVLTVAAPNSALLARALRAEPLIVGL
		1	i			ALAHSCLNPMLFLYFGRAOLRRSLPAACHW
		l			1	ALRESQGQDESVDSKKSTSHDLVSEMEV
851	2201	A	7011	1	2310	AAASPLRMSRKGPRAEVCADCSAPDPGWASI
	}					SRGVLVCDECCSVHRSLGRHISIVKHLRHSA
1	1	i			}	WPPTLLQMVHTLASNGANSIWEHSLLDPAQV
1		ł				QSGPALKQTPKDKV\HPIKSEFIRAKYQMLAF
1						VHKLPCRDDDGVTAKDLSKQLHSSVRTGNLE
						TCLRLLSLGAQANFFHPEKGTTPLHVAAKAG
		1				QTLQAELLVVYGADPGSPDVNGRTPIDYARQ
1		l	1.			AGHHELAERLVECOYELTDRLAFYLCGRKPD
		ļ				HKNGHYIIPQMADSLDLSELAKAAKKKLOAL
		İ				SNRLFEELAMDVYDEVDRRENDAVWLATON
						HSTLVTERSAVPFLPVNPEYSATRNQGROKL
		1	J			ARFNAREFATLIIDILSEAKRROOGKSLSSPTD
						NLELSLRSQSDLDDQHDYDSVASDEDTDQEP
		1				LRSTGATRSNRARSMDSSDLSDGAVTLOEYL
1			[]			ELKKALATSEAKVQQLMKVNSSLSDELRRLO
		1		•		REIHKLQAENLQLRQPPGPVPTPPLPSERAEH
		1	1			TPMAPGGSTHRRDRQAFSMYEPGSALKPFGG
1		l				PPGDELTTRLQPFHSTELEDDAIYSVHVPAGL
						YRIRKGVSASAVPFTPSSPLLSCSQEGSRHTSK
						LSRHGSGADSDYENTQSGDPLLGLEGKRFLE
1						LGKEEDFHPELESLDGDLDPGLPSTEDVILKT
İ			1			EQVTKNIQELLRAAQEFKHDSFVPCSEKIHLA
] .		l				VTEMASLFPKRPALEPVRSSLRLLNASAYRLQ
1						SECRKTVPPEPGAPVDFQLLTQQVIQCAYDIA
					_	KAAKQLVTITTREKKQ
852	2202	A	7016	484	1777	RISKIQVYYSTGYSSRKMNPTLGLAIFLAVLL
				ļ		TVKGLLKPSFSPRNYKALSEVQGWKQRMAA
						KELARQNMDLGFKLLKKLAFYNPGRNIFLSP
		1				LSISTAFSMLCLGAQDSTLDEIKQGFNFRKMP
					,	EKDLHEGFHYIIHELTQKTQDLKLSIGNTLFID
		-			1	-QRLQPQRKFLEDAKNFYSAETILTNFQNLEM
						AQKQINDFI/ESKTHGKINNLIENIDPGTVMLL
			[. [ĺ	ANYIFFRARWKHEFDPNVTKEEDFFLEKNSS
'					l	VKVPMMFRSGIYQVGYDDKLSCTILEIPYQK
				1	.	NITAIFILPDEGKLKHLEKGLQVDIFSRWKTL
					ļ	LSRRVVDVSVPRLHMTGTFDLKKTLSYIGVS
				ļ		KIFEEHGDLTKIAPHRSLKVGEAVNKAELKM
						DERGTEGAAGTGAQTLPMETPLVVKIDKPYL
]			LLIYSEKIPSVLFLGKIVNPIGK
853	2203	A	7017	1	3293	MTHACNPSTLGGQGRRITRSHGRRRSSRGPV
					ĺ	ARHVAAGAGHENKHGGSRRFPAGVAPRRAM
1					ł	ANVSKKVSWSGRDRDDEEAAPLLRRTARPG
						GGTPLLNGAGPGAARQSPRSALFRVGHMSSV
				ļ	1	ELDDELLEP\DMDPPHPFPKEIPHNEKLLSLKY
					I	ESLDYDNSENQLFLEEERRINHTAFRTVEIKR
]						WVICALIGILTGLVACFIDIVVENLAGLKYRVI
L						KGSILPNIDKFTEKGGLSFSLLLWATLNAAFV

0000 000	I one ve	132	1.000	1 20		
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in in	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ł	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ſ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	•	/=possible nucleotide deletion, \-possible
			<u> </u>	sequence	J	nucleotide insertion
						LVGSVIVAFIEPVAAGSGIPQIKCFLNGVKIPH
		1	ļ	ļ .		VVRLKTLVIKVSGVILSVVGGLAVGKEGPMI
	i	l				HSGSVIAAGISQGRSTSLKRDFKIFEYFRRDTE
						KRDFVSAGAAAGVSAAFGAPVGGVLFSLEEG
	İ					ASFWNQFLTWRIFFASMISTFTLNFVLSIYHG
						NMWDLSSPGLINFGRFDSEKMAYTIHEIPVFI
		1				AMGVVGGVLGAVFNALNYWLTMFRIRYIHR
		1				PCLQVIEAVLVAAVTATVAFVLIYSSRDCQPL
						QGGSM\$YPLQLFCADGEYNSMAAAFFNTPEK
	1	ļ				SVVSLFHDPPGSYNPLTLGLFTLVYFFLACWT
		1				YGLTVSAGVFIPSLLIGAAWGRLFGISLSYLTG AAIWADPGKYALMGAAAQLGGIVRMTLSLT
	ł					VIMMEATSNYTYGFPIMLVLMTAKIVGDVFIE
i	i		i			GLYDMHIQLQSVPFLHWEAPVTSHSLTAREV
		İ				MSTPVTCLRRREKVGVIVDVLSDTASNHNGF
	[İ				PVVEHADDTQPARLQGLILRSQLIVLLKHKVF
1						VERSNLGLVQRRLRLKDFRDAYPRFPPIOSIH
	[[VSQDERECTMDLSEFMNPSPYTVPQEASLPR
		l				VFKLFRALGLRHLVVVDNRNQVVGLVTRKD
	į	ł	1			LARYRLGKRGLEELSLAQTGPKAQATAEGRV
						AGAAQQPCQLRAVTLEDLGLLLAGGLASPEP
						LSLEELSERYESSHPTSTASVPEQDTAKHWNQ
			'			LEQWVVELQAEVACLREHKQRCERATRSLL
			1			RELLQVRARVQLQGSELRQLQQEARPAAQAP
						EKEAPEFSGLQNQMQALDKRLVEVREALTRL
						RRRQVQQEAERRGAEQEAGLRLAKLTDLLQ
				1		QEEQGREVACGALQKNQEDSSRRVDLEVAR M
854	2204	A	7037	139	2604	AGTWEPRPYDOAKETGAPGSOPPVPPMELRP
		••	1 '03'	137	2001	WLLWVVAATGTLVLLAADAQGOKVFTNTW
						AVRIPGGPAVANSVARKHGFLNLGQIFGDYY
						HFWHRGVTKRSLSPHRPRHSRLQREPQVQWL
				J		EQQVAKRRTKRDVYQEPTDPKFPQQWYL\SG
						VTQ\RDLMVKAAWAQGYTGHGIVVSILDDGI
	•					EKNHPDLAGNYDPGASFDVNDQDPDPQPRY
	•					TQMNDNRHGTRCAGEVAAVANNGVCGVGV
				Ì		AYNARIGGVRMLDGEVTDAVEARSLGLNPN
				l		HIHIYSASWGPEDDGKTVDGPARLAEEAFFR
				i		GVSQGRGGLGSIFVWASGNGGREHDSCNCD
			[[ļ		GYTNSIYTLSISSATQFGNVPWYSEACSSTLA
		1]		TTYSSGNQNEKQIVTTDLRQKCTESHTGTSAS
					}	APLAAGIIALTLEANKNLTWRDMQHLVVQTS
j					ł	KPAHLNANDWATNGVGRKVSHSYGYGLLD
			ŀ	-	ſ	AGAMVALAQNWTTVAPQRKCIIDILTEPKDI GKRLEVRKTVTACLGEPNHITRLEHAOARLT
		i	· I			
			}	- 1		LSYNRRGDLAIHI VSPMGTRSTI I AADDUDSU
						LSYNRRGDLAIHLVSPMGTRSTLLAARPHDY SADGFNDWAFMTTHSWDEDPSGEWVLEIEN
						SADGFNDWAFMTTHSWDEDPSGEWVLEIEN
				1		SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS
						SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP
						SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS
	,]					SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS
						SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS
				,		SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS
						SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD
855	2205	A	7058	3		SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL
855	2205	A	7058	3	. 1441	SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL QRPASQLLAPFAAEALPGAPRAAMAQHFSLA
855	2205	A	7058	3	. 1441	SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFVTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL
855	2205	A	7058	3	. 1441	SADGFNDWAFMTTHSWDEDPSGEWVLEIEN TSEANNYGTLTKFTLVLYGTAPEGLPVPPESS GCKTLTSSQACVVCEEGFSLHQKSCVQHCPP GFAPQVLDTHYSTENDVETIRASVCAPCHAS CATCQGPALTDCLSCPSHASLDPVEQTCSRQS QSSRESPPQQQPPRLPPEVEAGQRLRAGLLPS HLPEVVAGLSCAFIVLVFYTVFLVLQLRSGFS FRGVKVYTMDRGLISYKGLPPEAWQEECPSD SEEDEGRGERTAFIKDQSAL QRPASQLLAPFAAEALPGAPRAAMAQHFSLA ACDVVGFDLDHTLCRYNLPESAPLIYNSFAQF

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YFDLPGALLCARVVDYLTKLNNGQKTFDFW KDIVAAIQHNYKMSAFKENCGIYFPEIKRDPG RYLHSRPESVKKWLRQLKNAGKILLLITSSHS DYCRLLCA\YILGNDFTDLFDIVITNALKPGFP SHLPSQRPFRTLENDEEQEAI.PSLDKPGWYSQ GNAVHLYELLKKMIGKPEPKVVYFGDSMHS DIFPARHYSNWETVLIEELRGDEGTRSQRPE ESEPLEKKGKYEGPKAKPLNTSSKKWGSFF\I DSVLGLENTEDSLVYTWSCKRISTYSTIAIPSI EAIAELPLDYKFTRFSSSNSKTAGYYPNPPLV
856	2206	A .	7082	396	1635	LSSDETLISK SSPSVFEFEHAVQPVFTMEFLKTCVLRRNACT AVCFWRSKVVQKPSVRRISTTSPRSTVMPAW VIDKYGKNEVLRFTQNMMMPIHYPNEVIVK VHAASVNPIDVNMRSGYGATALNMKRDPLH VKIKGEEFPLTLGRDVSGVVMECGLDVKYFK PGDEVWAAVPPWKQGTLSEFVVVSGNEVSH KPKSLTHTQAASLPYVALTAWSAINKVGGLN DKNCTGKRVLILGASGGVGTFAIQVMKAWD AHVTAVCSQDASELVRKLGADDVIDYKSGSV EEQLKSLKPFDFILDNVGGSTETWAPDFLKK WSGATYVTLVTPFLLNMDRLGIADGMLQTG VTVGSKALKHFWKGVHYRWAFFMASGPCL DDIAELVDAGKIRPVIEQTFPFSKVPEAFLKV ERGHARGKTVINVV
857	2207	A	7088	320	2417	LRRKMTPQSLLQTTLFLLSLLFLVQGAHGR GHREDFRPCSQRNQTHRSSLHYKPTPDLRISIE NSEEALTVHAPFPAAHPASRSFPDPRGLYHFC LYWNRHAGRLHLLYGKRDFLLSDKASSLLCF QHQEESLAQGPPLLATSVTSWWSPQNISLPSA ASFTFSFHSPPHTGAHNASVDMCELKRDLQL LSQFLKHPQKASRRPSAAPASQLQSLESKLT SVRFMGDMGSFEEDRINATVWKLQPTAGLQ DLHHSRQEEEQSEIMEYSVLLPRTLFQRTKG RSGEAEKRLLLVDFSSQALFQDKNSSQVLGE KVLGIVVQNTKVANLTEPVVLTFQHQLQPKN VTLQCVFWVEDPTLSSPGHWSSAGCETVRRE TQTSCFCNHLTYFAVLMVSSVEVDAVHKHY LSLLSYVGCVVSALACLVTIAAYLCSRVPLPC RRKPRDYTIKVHMNLLLAVFLLDTSFLLSEPV ALTGSEAGCRASAIFLHFSLLTCLSWMGLEG YNLYRLVVEVFGTYVPGYLLKLSAMGWGFPI FLVTLVALVDVDNYGPIILAVHRTPEGVIYPS MCWIRDSLVSYITNLGLFSLVTLFNMAMLAT MVVQILRLRPHTQKWSHVLTLLCLSLVLGLP WALIFFSFAGTFQLVVLYLFSITSFQGFLIFI WYWSMRLQARGGPSPLKSNSDSARLPISSGS TSSSRI
858	2208	A	7091	185	415	DAGAVKSSDTNIWFRGMCDDKKGHRCPS*G QPQHFHVAFHTEAEGAMFYFRLHVIHRVMQS QQQLFPSTLFSWLLE
859	2209	A	7136	3	302	FFFWRQSLALLPRLECSGATGAHCNLHFPGSS DCPTSAS*IAGITGACYHAWLLFVFLAETGFH HVGQGGLELLTSSDPSGSASQSAGITGVSHCT WPI
860	2210	A	7156	23	591	ALSTETRTPDMRRLLLVTSLVVVLLWEAGAV PAPKVPIKMQVKHWPSEQDPEKAWGARVVE PPEKDDQLVVLFPVQKPKLLTTEEKPRGQGR GPILPGTKAWMETEDTLGRVLSPEPDHDSLY

SEQ ID NO: of nucl-	SEQ ID NO: of peptide	Met hod	SEQ ID NO: in	Predicted beginning nucleotide	Predicted end nucleotide location	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
cotide seq- uence	seq- uence		USSN 09/496 914	location correspondi ng to first amino acid	to last amino acid residue of peptide	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan,
				residue of peptide sequence	sequence	Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion HPPPEEDQGEERPRLWVMPNHOVLLGPEEDQ
					 	DHIYHPQ*GSRGHHCPRPVPRPRLLGLGPSLP CPS
861	2211	A	7161	1220	1003	NYVCTIAF*EKKMGF*LSLSCLVLLFVLFLDCI LTTTTRIMFHCTYLFASVCLSLLNTLLSPNCL KSAMILQ
862	2212	A	7211	665	847	LKYYHITMGIYKTGKKVIL*KSSMSNRFSVIF YKNIQKLSFSNYVYHQNYVFSSDWSYDF
863	2213	A	7212	924	1273	HGSSCALGDLAPG*LPSGPVLSSPAVRL*RKP LVWDSPSCLPATGPT*GLVLVLGGPDCT*WA RGQHEHKRMRAP*SCRVTVNLAKKKKKTDQ CIKPNYQSPPKECDYNILANSVA
864	2214	A	7214		1619	SDKGGKKADRKNHLRHAFPLLPHRVRERLH DPKVPVDADHVQGQDPGRAAHDIHGEDVTE KVSKDPLAPDEVGDTDEGHDRHGHREVGQR HGHDQEEVAYEERACEGGKFATVEVTDKPV DEALREAMPKVAKYAGGTNDKGIGMGMTV PISFAVFPNEDGSLQKKLKVWFRIPNQFQSDP PAPSDKSVKIEEREGITVYSMQFGGYAKEAD YVAQATRLRAALEGTATYRGDIYFCTGYDPP MKPYGRRNEIWLLKT
865	2215	A	7246	559	682	RRLGAVAHAYTSSTLGGRGGWIT*GQELQTS LANMAKPRLY
866	2216	A	7257	641	1310	TCTYKYLMGWIRGRRSRHSWEMSEFHNYNL DLKKSDFSTRWQKQRCPVVKSKCRENASPFF FCCFIAVAMGIRFIIMVAIWSAVFLNSLFNQEV QIPLTESYCGPCPKNWICYKNNCYQFFDESKN WYESQASCMSQNASLLKVYSKEDQDLLKLV KSYHWMGLVHIPTNGSWQWEDGSILSPNLLT IIEMQKGDCALYASSFKGYIENCSTPNTYICM QRTV
867	2217	A	7288	151	396	SIKIIEAFGSNGPDFWFFRYWSP*LFRQQVVFI MPFFQTLWLMNANRFCSIFTTTNVANNCWW TPYHCWLSVVVCRCESHGI
868	2218	A	7298	3	272	PDTVIGGRGSGGKEFGRWVLW*VFE*RLGTP KGSCPAGGSRMVSESD*EGRGC*ASYPCAC* AGS*WR*GSRPAGRGTPPRSLSHARPP
869	2219	Α	7332	1223	332	PRRDAEDRDESCLNPAFPIGLLHPNSVNSMAR FLTLCTWLLLLGPGLLATVRAECSQDCATCS YRLVRPADINFLACVMECEGKLPSLKIWETC KELLQLSKPELPQDGTSTLRENSKPEESHLLA KRYGGFMKRYGGFMKKMDELYPMEPEEEA NGSEILAKRYGGFMKKDAEEDDSLANSSDLL KELLFTGDNRERSHHQDGSDNEEEVSKRYGG FMRGLKRSPQLKEKAKELQKRYGGFMRRVG PQKW*MTSPQNRYGGFLKRFAEALPSDEEGE SYSKEVPEMEKRYGGFMRF
870	2220	A	7382	216	1018	EIHQRLTERTQFLDESRKNPNS*QANLLRGGG AGQGRGREGAESGGSRGFGPGSDGRLPATGD FWSPRSQRRGCCGRRAPRPEAMENGAVYSPT TEEDPGPARGPRSGLAAYFFMGRLPLLRRVL KGLQLLLSLLAFICEEVVSQCTLCGGLYFFEF VSCSAFLLSLLILIVYCTPFYERVDTTKVKSSD FYITLGTGCVFLLASIJFVSTHDRTSAEIAAIVF GFIASFMFLLDFITMLYEKRQESQLRKPENTT RAEALTEPLNA
871	2221	A	7403	3	393	SCAMCSGLL*LLLPIWLSWTLGTRGSEPRSVN DPGNMSFVKETVDKLLTGFRCFREREAAPRR ALRGAALPGESEAGDPESLRSSVNADWIQYS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ļ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ŀ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
]		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
Į	,	,		sequence		nucleotide insertion
				004		DLWEAEVSTPRCEAGFCQECFRTPGNQEKDG
		ļ				PFIC
872	2222	A	7413	1061	359	FVDIVSVVEFPHCPEARFPAQHGQDSKRLTLC
0,2		l "	'*15	1001	337	PGGS*PQATLHLDRMRVSASPTKEIQVKKYK
						CGLIKPCPANYFAFKICSGAANVVGPTMCFED
ł		ł			ł	RMIMSPVKNNVGRGLNIALVNGTTGAVLGO
		ì				KAFDMYSGDVMHLVKFLKEIPGGALVLVAS
			ļ			YDDPGTKMNDESRKLFSDLGSSYAKQLGFRD
		Į				SWVFIGAKDLRGKSPFEQFLKEQPQTQNKYE
		l	ĺ	!	ĺ	GWPELLEMEGCMPPKPF
873	2223	A	7429	2242	2394	ILKCAGHGGSCL*SQHFGRLRWEDRLRLGVO
0/3	2223	^	7423	2242	2374	
874	2224	_	7460	146	894	DHPGQHCETPSLLKIERKLF
0/4	2224	Α	7468	146	074	PCTSCVLWATLHLPASTRKAPQAECGMISITE
1			1			WQKIGVGITGFGIFFILFGTLLYFDSVLLAFGN
1		Ì				LLFLTGLSLIIGLRKTFWFFFQRHKLKGTSFLL
l .		1)	GGVVIVLLRWPLLGMFLETYGFFSLFKGFFPV
ļ		ļ				AFGFLGNVCNIPFLGALFRRLQGTSSMV*KTE
'		1	i .		İ	MSSLNLDHWLKGAKREEWEPPPQSPALTHSP
1		i	•		ł	TYPGPPQVQKERNGAEQLTSNPQVDSRGCQE
		<u> </u>				AEMQTPRRLGWGWYHTLTLYLWEEK
875	2225	Α	7498	91	251	GEKPVPTWLQDEAGQWLLGFVAQPWGWPG
						SERHEP*HGGVLFRLGPSAPPGKL
876	2226	Α	7544	403	587	YSCLCFLFKHITSFKNSVHIWLGTVVHAYNPN
						ILGGQGGWIA*GQEFKTSLGNTVRPCLYK
877	2227	A	7566	2	940	GCAPDTRFFVPEPGGRGAAPWVALVARGGC
1 .		ļ]	TFKDKVLVAARRNASAVVLYNEERYGNITLP
						MSHAGTGNIVVIMISYPKGREILELVQKGIPV
		l				TMTIGVGTRHVQEFISGQSVVFVAIAFITMMII
		i		•		SLAWLIFYYIQRFLYTGSQIGSQSHRKETKKVI
		l			1	GQLLLHTVKHGEKGIDVDAENCAVCIENFKV
		ł				KDIIRILPCKHIFHRICIDPWLLDHRTCPMCKL
]	1			DVIKALGYWGEPGDVQEMPAPESPPGRDPAA
,,					,	NLSLALPDDDGSDESSPPSASPAESEPQCDPSF
Ĺ		l				KGDAGENTALLEAGRSDSRHGGPIS
878	2228	Α	7586	315	1232	ERSLLCKVDVRWIYVSEGTKTQRRHRQGSLR
		l	Ì			RGRMQAACWYVLFLLQPTVYLVTCANLTNG
		ŀ				GKSELLKSGSSKSTLKHIWTESSKDLSISRLLS
		ſ				QTFRGKENDTDLDLRYDTPEPYSEQDLWDW
		1				LRNSTDLQEPRPRAKRRPIVKTGKFKKMFGW
		1	1			GDFHSNIKTVKLNLLITGKIVDHGNGTFSVYF
1		ļ				RHNSTGQGNVSVSLVPPTKIVEFDLAQQTVID
		1				AKDSKSFNCRIEYEKVDKATKNTLCNYDPSK
		l				TCYQEQTQSHVSWLCSKPFKVICIYISFYSTD
	-	j				YKLVQKVCPDYNYHSDTPYFPSG
879	2229	Α	7605	479	391	TESWKLKWWSPTCLDQLNGSAPGNVFIHG
880	2230	A	7612	93	659	DAAVAMTAQGGLVANRGRRFKWAIELSGPG
		!]	GGSRGRSDRGSGQGDSLYPVGYLDKQVPDTS
	•	Ì				VQETDRILVEKRCWDIALGPLKQIPMNLFIMY
1		ľ	1	1		MAGNTISIFPTMMVCMMAWRPIOALMAISAT
		[FKMLESSSOKFLOGLVYLIGNLMGLALAVYK
		1				COSMGLLPTHASDWLAFIEPPERMEFSGGGL
						LL
881	2231	A	7615	291	1452	SPOKTMRSHTITMTTTSVSSWPYSSHRMRFIT
""		١.,	, , , ,		4700	NHSDQPPQNFSATPNVTTCPMDEKLLSTVLTT
		1				SYSVIFIVGLYGNIIALYVFLGIHRKRNSIQIYL
		1				LNVAIADLLLIFCLPFRIMYHINQNKWTLGVIL
		1				CKVVGTLFYMNMYISIILLGFISLDRYIKINRSI
		l				QQRKAITTKQSIYVCCIVWMLALGGFLTMIIL
		1				TLKKGGHNSTMCFHYRDKHNAKGEAIFNFIL
L	L		<u> </u>	L	I	ILAKUUTINSIMUTTI KUKHNAKUEAITNITU

SEQ ID NO: of much-cotide sequence (A-Alamine C-Cysteine, Designing) in unclostide peptide sequence (Markoviter) in unclostide location plant of the peptide sequence (Markoviter) in unclostide location may be unclosed to peptide sequence (Markoviter) in unclostide location in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of peptide sequence (Markoviter) in go to first a minio acid eric development of go to first a minio acid eric development of go to first a minio acid eric development of go to first a minio acid eric development of go to first a minio acid eric development of go to first a minio acid eric development of go to first a minio acid eric development of go to first a minio acid eric development of go to first a minio acid eric development of go to first a minio acid eric development of go to first a minio acid eric development of go to first a minio acid eric development of go to first a minio acid eric development of go to first a						 	
mucle citide ssp. uence USSN location go fairs location go fairs location go fairs location go fairs location go fairs location go fairs location go fairs location go fairs location go fairs location go fairs location go fairs location go fairs location go fairs location go fairs location locat	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Section	1		hod				
		1					
1			İ				
### ### ### ### ### ### ### ### ### ##	1 -	uence					
Pesidue of peptide pep	uence	!	1	914			
Peptide Pept		1					
		1]			seduence	
			1				
SGKYATTARNSPIVLIEFTICFUPYHARRIPITISE QLNVSSCYWEEVIKTEMEM_USSPISCI_D	<u> </u>	 -	\vdash	 	sequence		
						}	
WMYFLMSSNINKIMCQLLFRRFQGEFSRSEST	i	ł	ł	1	ł		· ·
SEFKPGYSLHDTSVAVIGIOSSKST	1		l				
882 2232 A 7617 67 379 ROMALIKANKDLISAGIKEFSVLLINQOVPOTO PLYSEEDMYTVYDEDWMSTYDTYYRQOYTGE POERDKALOELROELMTLANPFLAKYRDFLK SHELPSHIPTSS SHELPSHIPTSS STATEMENT SHELPSHIPTS STATEMENT SHELPSHIPTS STATEMENT SHELPSHIPTS STATEMENT SHELPSHIPTS STATEMENT SHELPSHIPTS STATEMENT SHELPSHIPTS STATEMENT SHELPSHIPTS STATEMENT SHELPSHIPTS STATEMENT SHELPSHIPTS STATEMENT SHELPSHIPTS SHELPSHIPTS STATEMENT SHELPSHIPTS SHELPSH	1			1			
PL-VSEEDMVTVVEDWMS-Y19TY-QG-VTGE POERDKALQEL ROLL-NTLANFFLAKYRDELK SHELPSHIPPSS	882	2232	A	7617	67	379	
POERDKALQELRQELNTLANPFLAKYRDFLK	1002	22,2	^^	/**/] "	3.,	, ,
SHELPSHPPTS 883 2233 A 7622 400 215 XVKTCRYPRKYSAANDTGFVDIPSREKDLAK AVATVGPISVAVGASHVFFOFYKKGKHLSS 884 2234 A 7638 2640 2861 APVILIQMVKLSIVLTPQT-SHIDQGQTRELING QHVKXSTQFCPTJRKVSECQMGFGALEQFP GLSCHTSHSG GLSCHTSHSG 885 2235 A 7642 201 455 PSRGKMELEAMSRYTSPVNPAVFPHLTVVLL AIGMFTTAWFFVYEVTSIKYTRDIVKELLISL VASLEMGFOUL-FLLWGIFYV 886 2236 A 7692 61 569 APENFFSRQIFNSETKVKLSLKTGTWLGNHA HLGHFSTHHELGLSGKVVGPLVKNILEVIRN GGMETRHPGKVSSWFIR WBSRAEQHNHAE HHEDVPQGDEDSKVSEAQQEFPDVVTCAGL GLIPKALRVLLFQLKVQHRRGHQRFPQQD VSDHRYGRSVRQNRK 887 2237 A 7693 85 315 NPGCLPVAMRTSYLLFTLCLLSEMASGG NFLTGLGHRSDHYNCVSSGGCL YSACPIFTK IGGTCYRGKAKCK GGMETRHPGKVSSWFIR WBSRAEQHNHAE HHEDVPQGDEDSKVSEAQQEFPDVVTCAGL GURRARVSTRYSLLFTLEXVQHRRGHQRFPQQD VSDHRYGRSVRQNRK VSDHRYGRSVRQNRK R88 2238 A 7702 242 1298 APSHRRRYLSPSRSAGQLGMALERLCSVLK VLLITVLVVEGLAVAQKTQDQQNIGIKHIPA QCGWRTSNGGHASPHYDENYPRKECTYI LEAAPRQRELITPDEHYYIERSFECRFDHLEVY DGPGFSPLDRYGVKSSPLISTGRFMWIKF SSDEELEGLGFRAKYSFIPDDFTYLGGLMPIP CQFGRSTNAKSVGVEGEKKFTGQAVD CWTIKATFKAKIVLRFLDYQMEHSNECKRY VAVYOGSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSSEKKTPGQAVD CRYMINDSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLMPDDLREEFVKAGLKCGP TISTTRFFFKKLAQALLEQGGRLSSFYHHEA GVTALSOPPORILKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ATASSEKTAPILAVOG GCRTNYNHIVAKAKNQASICQHTOXPLENINYV NKEALQAVKMIKGSBRKAFSTREDAEKFAR GICYFPSPSPSKTSLPLSPVKTAPLFSNDLAFKAR GCRTNYNHIVAKAKNQASICQHTOXPLENINY NKEALQAVKMIKGSBRKAFSTREDAEKFAR GICYFPSPSPSKTSLPLSPVKTAPLSPNDLAFKAR GLGRSSTTVNKRRANSYKNRTQDLTAKKR VKNSRNKVOKTPEDVICERSKNSSVELERF EVLKGHYYPULLAREFTSSPVGBL WSPDOTA EASHVSRYGGSPDDPVLTLRAFAGPLSPAKAE DFRKLWKTPREKAGFLHVYKKSDFERGFE VKGHAHELSPLAYGGSPRDF GGRTTHTUR VKNSRNKVOKTPEDVICERSKNSSVELERF VKGHAHELSPLAYAGE GGRTTHTUR VKNSRNKVOKTPEDVICERSKNSSVELERF	1		1				
883 2233 A 7622 400 215 KVKTCRYNPKYSAANDTGPVDIPSREKDLAK AVATVGPISVAVQASHVPEPCKGKHLSS		ľ		ì			
AVATVOPISVAVGASHVFFQFYKKGKHLSS	883	2233	A	7622	400	215	
884 2234 A 7638 2640 2861 APVILIQMYKLSIVLTPQFLSHDQGQLTKELQ QHVKSVTCPCEYLRK VSECRQMGPGALEQFP GLSCHTSHSG 885 2235 A 7642 201 455 PSRGKMELBAMSRYTSPVNPAVPFHLTVVLL ALGMFTA WFVYEVTSTKYTYKDIVKELLISL VASLFMGFGVLFLLLWGIVV ALGMFTA WFVYEVTSTKYTYKDIVKELLISL VASLFMGFGVLFLLLWGIVV ALGMFTA WFVYEVTSTKYTYKDIVKELLISL VASLFMGFGVLFLLLWGIVV ALGMFTA WHERDYPGGEDSKYSEAQGEPVVTCAGLP GLIPKALRVLLFQLKVQHRRGIHQQRPEQQD VSDRYGRONGK A 7692 61 569 APENPFSRQHFNSETKVKLSLKTGTWLGNHA HLGEHFSTHHELGLSGKVVGFLVKNILEVIRN GGMETRHPGKVSSWFHRWDSRAEQHNHAE HLBDVPGGEDSKSVSEAQGEPVVTCAGLP GLIPKALRVLLFQLKVQHRRGHQRPEQQD VSDRYGRONGK 887 2237 A 7693 85 315 NPGCCLPVAMRTSYLLLFTLCLLSEMASGG NFLTGGHRSDHYNCVSSGGQCLYSACPIFTK IQGTCYRGKAKCK RSDRYGGRASPHYNCVSSGGQCLYSACPIFTK IQGTCYRGKACKCK 4 APSHRRRYLSPSRSAGGLGNMALERLCSVLK VLLITVLVSGLAVAQKTOGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTDEHTYIEPSSECREPHLEVR QCGIWVRTSNGGHRASPNYPDSYPPNKECIYI LEAAPRQRIELTDEHTYIEPSSECREPHLEVR QCGIMPIP QCGPGFSSLDAYGKCGYRSSQVEQEKTKPGQAVD CIWIKATPKAKIYLRFLDVQMEHSNECKRNF VAYVDGSSIENLKAKRRITTSGGASPAQA ALSFCHSMCNISKAKRSTVANDVMLKTGI GVIRMWADGGSRLNAFRAHTISTGGASPAQA ALSFCHSMCNISKAKRSTVANDVMLKTGI GVIRMWADGGSRLNAFRAHTISTSGGASPAQA ALSFCHSMCNISKAKRSTVHAGKCOR FORTAM QCGIWVRTSNGGSILGLFGFFGLGLG MTMTMALLALLALLALLALLALLALLALLALLALLALLALLALL			1				•
September Sept	884	2234	A	7638	2640	2861	
		'	1	1			
886 2236 A 7692 61 569 APENPFSRQHFNSETKVRLISLVVGLVSKULSLSLS 886 2236 A 7692 61 569 APENPFSRQHFNSETKVRLSLKTGTWLGNHA HLGEHFSTHHELGLSGKVVGFLVSKULEVIRN GGMETRHFGKVSSWFHARWDSRAEQHNHAE HHEDVPQGDEDSKVSEAQQEFPDVVTCAGLP GLLPKALRVLLFQLKVQHRPGHQQRPEQQD VSDHRYGRSVVQRVRK 887 2237 A 7693 85 315 NPGCCLPVANRTSYLLLFTLCLLISMASGG NFLTGLGHRSDHTNCVSSGQCLYSACPIFTK IQGTCYRGKAKCCK 888 2238 A 7702 242 1298 APFRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGLAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPNECTYI LEAAPRQRIELTFDEHYYLIFSFECRFDHLEVR DGPFGSPLDRYGGVSSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSQVGQEEKTKPQQAVD CIWTIKATFKAKIYLR-LDVQMEHSDECKNFY VAVYDGSSISENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYMNPSTHIPASAGGSILGLFDFFGLGLGE MTMDALLARLKILNPDDLREIVKAGLKCGP ITSTTRIFEKKLAQALLEQGGRLSSPYHEBA GVYASGNPPEEBAVTISKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDYPARNETIYVYE NKKBALQAKMIKGSRFKAFSTREDAEKFAR GICDYPPSPSKTSLPJSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTTSDLIWSNRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLITDVLENP DFMRLMYPDDDEAMLGKRRYVVDLLYINTD DKMGYDTPLHFACKFGRADVDNVLSSHHLI VKNSRNKYDKTPETCDVLCRSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDVLTLRAFAGGPER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD FORKLWKTPPREKAGFLHVKSDPERGGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD	ļ					ļ	
886 2236 A 7692 61 569 APENPFSRQHFNSETKVRLISLVVGLVSKULSLSLS 886 2236 A 7692 61 569 APENPFSRQHFNSETKVRLSLKTGTWLGNHA HLGEHFSTHHELGLSGKVVGFLVSKULEVIRN GGMETRHFGKVSSWFHARWDSRAEQHNHAE HHEDVPQGDEDSKVSEAQQEFPDVVTCAGLP GLLPKALRVLLFQLKVQHRPGHQQRPEQQD VSDHRYGRSVVQRVRK 887 2237 A 7693 85 315 NPGCCLPVANRTSYLLLFTLCLLISMASGG NFLTGLGHRSDHTNCVSSGQCLYSACPIFTK IQGTCYRGKAKCCK 888 2238 A 7702 242 1298 APFRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGLAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPNECTYI LEAAPRQRIELTFDEHYYLIFSFECRFDHLEVR DGPFGSPLDRYGGVSSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSQVGQEEKTKPQQAVD CIWTIKATFKAKIYLR-LDVQMEHSDECKNFY VAVYDGSSISENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYMNPSTHIPASAGGSILGLFDFFGLGLGE MTMDALLARLKILNPDDLREIVKAGLKCGP ITSTTRIFEKKLAQALLEQGGRLSSPYHEBA GVYASGNPPEEBAVTISKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDYPARNETIYVYE NKKBALQAKMIKGSRFKAFSTREDAEKFAR GICDYPPSPSKTSLPJSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTTSDLIWSNRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLITDVLENP DFMRLMYPDDDEAMLGKRRYVVDLLYINTD DKMGYDTPLHFACKFGRADVDNVLSSHHLI VKNSRNKYDKTPETCDVLCRSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDVLTLRAFAGGPER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD FORKLWKTPPREKAGFLHVKSDPERGGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD	885	2235	A	7642	201	455	
		"	1				
886 2236 A 7692 61 569 APENPFSRQIFINSETKVKLSLKTGTWLGNIH. HLGEHFSTHHELGLSGKVVGFLVKNILEVIRN GGMETRHPGKVSSWFHRWDSRAEQHNIHAE HHEDVPQGBEDSKVSEAQQEFPDVVTCAGLP GLPKALVLLFQLKVQHRPGIHQQRPEQQD VSDHRYGRSVRQNRK 887 2237 A 7693 85 315 NPGCCLPVAMRTSYLLLFTLCLISEMASGG NFLTGLGHRSDHYNCVSSGGQCLLSEMASGG NFLTGLGHRSDHYNCVSSGGQCLYSACPIFTK IQGTCYRGKAKCCK 888 2238 A 7702 242 1298 APSHRRRYLSPSRAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DCPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPPDFTYLGGILNPIP DCQFELSGADGVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKNIP VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNKAFSTSVPSPTVLGRLNCP ITSTTRFFEKKLAQALLEQGGRLSSFYHBEA GVTALSQDPQRILKPAEGNFTDQAGFSEDRDF GYSVGLNPPEEAVTSKTCSVPYPSDTDTYRAG ATASKEPPLYYGVCVBVPARNERITYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFFSPSKTSLPLSFVKTALFSNDRLKDG LCLSESSTVNKERANSYKNPRTDDLTAKLKK AVEKGEEDTFSDLIWSNPRVLIGSGDNPTIVQ EGKRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLOKRIRVYVLLVLNTIP DKMGYDTPLHFACKFONADVNVLSSHHLI VKNSRNKVDKTPEDVLAKGLKGRR EYLKGHYYVFLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSFRDPVLTLRAFGDLSNSCHE,KERR EYLKGHYYVFLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSFRDPVLTLRAFGENSKNSVELKERR EYLKGHYYVFLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSFRDPVLTLRAFGEPLSPAKAA FYLKGHYYVFLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSFRDPVLTLRAFGEPLSPAKAA FYLKGHYYVFLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSFRDPVLTLRAFGEPLSPAKAA FYLKGHYYVFLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSFRDPVLTLRAFAGELSPAKAC FYLKWTTPREKAGGHHHVKKSDPERGFGFR VGRELAHELGYWVEWEFLGCCFVDLSSQE GLGRLEEYLTQQEIGKKAQQETTGEREASCRD KATTSGSNSISVARTLDEDDMSLEEIKNRQNA				1	1		
887 2237 A 7693 85 315 NPGCCLPVAMRTSTLLLSEMASGG NPLTGLGHRSDHYNCVSSGQCLYSACPIFTK IQGTCYRGKAKCCK 888 2238 A 7702 242 1298 APSHRRYLSPSKAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDQNIGIKHIPAT QCGIWVRTSNGGHFASSNYDDSYPPKECIYI LEAAPRQRIELTFDEHYYIEPSFECKPDHLEVR DGPFGFSPLIDRYCQVKSPPLIRSTGRFMWKF SSDELEGLGGRAKYSFIPDPTTYLGGILNPIP DCQFELSGADGRAKYSFIPDPTTYLGGILNPIP DCQFELSGADGRAKYSFIPDYTQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLINGRAMFSTENGASPAQA ALSFCHISNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREINKAGLKCGP TISTTRFIPEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEBAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYVGVCPYYEDVPARNERTYVYE NKKEALQAVKMIKGSRKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNERRANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPITVQ EGCRYNVMHVAAKROASICQLITDVLENP DFMRLMYPDDDEAMLQKRRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNLSSHHLI VNNSRNKYDKTPEDVICERSKNKSVELKERR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHYSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWWEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD	886	2236	A	7692	61	569	APENPFSRQHFNSETKVKLSLKTGTWLGNHA
887 2237 A 7693 85 315 NPGCCLPVANRTSYLLIFTLCLLSEMASGO NFLTGLGHRSDHYNCVSSGQCLYSACPIFTK IQGTCYROKAKCCK 888 2238 A 7702 242 1298 APSHRRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDQQNIGIKHIPAT QCGIWVRTSNGGHFASSNYDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECKFDHLEVR DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEKTKRFQQAYD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMPSTHHFASAGGSILGLFDFFGLGLE MTMDALLARIKLLHPDDLREEFWKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPEEBELNSTKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRKAKATSTREDAEKFAR GICDYFPSPSKTSLYSKTALSPSDPLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEDITSDLIWSNPRYLIGSGDNPTIVQ EGGRYNVMHVAAKENQASICQLILDYLENP DFMRLMYPDDDEAMLQKRIRYVVDLXLNTP DKMGYDTPLHFACKFGNADVVNULSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYVYPLLRAEETSSPVIGELWSPDOTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVARAEDDDDMSLEEIKNRQNA		}			<u></u>		HLGEHFSTHHELGLSGKVVGFLVKNILEVIRN
887 2237 A 7693 85 315 NFGCCLPVAMRTSYLLLFTLCLLLSEMASGG NFLTGLGHRSDHYNCVSSGQCLYSACPIFTK IQGTCYRGKAKCKER 888 2238 A 7702 242 1298 AFSHRRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDGQNIGIKHIBAT QCGIVYRTSNGGHFASPNYPDSYPNKECTYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DCFFGFSPLIDRYCGVKSPFLRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGIN.PIP DCQFELSGADGIVRSSQVEQEKTKPGQAVD CIWTIKATPKAKTYLRFLDYQMEHSNECKNIP VAVYDGSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNOQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMPSTHHPASAGGSILGI.FDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAGGNPTDQAGFSEDRDF GYSVGI.NPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRKARSTREDAEKFAR GICDVFPSPSKTSLPJ.SPVKTAPLFSNDRLKDG LCLSSETVNKERANSYKNPRTQDLTAKLRK AVEKGEDITSDLIWSNRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLI.DVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSHLIJ VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYVYPLLRAEETSSPVIGELWSPDOTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDFERGFER VGRELAHELGYPWWEYWEFLGCFVVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA		1	ļ	}			
887 2237 A 7693 85 315 NPGCCLPVAMRTSYLLLFTLCLLLSEMASGG NFLTGLGHRSDHYNCVSSGGQCLYSACPIFTK IQGTCYRGKAKCCK 888 2238 A 7702 242 1298 AFSHRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGLAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYEPSFECRFDHLEVR DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEEKTKPQQAVD CIWTIKATFKAKHYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCNNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYDROLDERLEVKAGLKCGP ITSTTRFFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAGGNIPDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYPPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSSESTVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGGRYNVMHVAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQRRRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHILI VKNSRNXYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPEKAGFLHHVKKSDPERGFER VGRELAHELGYYWVEYWEFLGGFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVSYRAFLDEDDMSLEEIKNRQNA		J	ļ	j]		HHEDVPQGDEDSKVSEAQQEFPDVVTCAGLP
887 2237 A 7693 85 315 NPGCCLPVAMRTSYLLLFTLCLLLSEMASGG NFLTGLGHRSDHYNCVSSGQCLYSACPIFTK 1QGTCYRGKAKCK 888 2238 A 7702 242 1298 APSHRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDQQNIGIKHIPAT QCGWVRTSNGHASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSPLIRSTGRFMWIKF SSDEELEGLGFRAKYSPPLIRSTGRFMWIKKTGI GVIRTMADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSSESTVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSIUWSNPRYLIGSGDNPTTVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRRYVVDLYLNTP DKMGYDTPLHFAKGFGNADVVNVLSSHILJ VKNSRNKYDKTPDDLTAKLRK AVEKGEEDTFSDVIGELMSPDQTTA EASHVSRYGGSRCKFGNADVVNVLSSHILJ VKNSRNKYDKTPDDVLTAKLRK AVEKGEEDTFSDVIGELSFSPVIGELWSPDQTA EASHVSRYGGSRDPVLTLRAFAGFLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYTWVEVWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	}			1		1	GLLPKALRVLLFQLKVQHRPGIHQQRPEQQD
888 2238 A 7702 242 1298 APRRPYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECFDHLEVR DGFFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEEKTKPQQAVD CIWTIKATPKAKKYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLINRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEVKAGLKCGP ITSTTRFIFEKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRLIKPASGNPTDQAGFSEDRDF GYSVGLNPPEEBAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYPPSPSKTSLPLSSPVTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVHHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYINTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNK VDKTPEDVICERSKNKSVELKERIR EYLKGHYVPLLRAEETSSPVIGELWSPDQTIA EASHVSRYGGSPRDPVLTLRAFAGFLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPEGFER VGRELAHELGYPWVEVWEFLGCFVDLSSQE GLQRLEFYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISTVAFLDEDDMSLEEIKNRQNA	L						
888 2238 A 7702 242 1298 APSHRRRYLSPSRSAGQLGNMALERLCSVLK ULITULVVEGIAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DGPFGFSPLIDRYCGVKSPFLRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEGEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP GYSVGLNPPEEBAVTSKTCSVPPSDTDTYRAG GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEBAVTSKTCSVPPSDTDTYRAG ATASKEPPLYQCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVERGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGFLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNISVRAFLDEDDMSLEEIKNRQNA	887	2237	A	7693	85	315	
888 2238 A 7702 242 1298 APSHRRRYLSPSRSAGQLGNMALERLCSVLK VLLITVLVVEGIAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DGPFGFSPLIDRYCGVKSPFLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKFAGNFTDQAGFSEDRDF GYSVGLNPPEEBAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPJSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMILMYPDDEAMLOKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYJKGHYYPPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPPWEYWEFLGCFVDLSSGE GLQRLEEYLTQGEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	1		ł	1	!		•
VLLITVLVVEGIAVAQKTQDGQNIGIKHIPAT QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEKKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNISVRAFLDEDDMSLEEIKNRQNA	<u> </u>	L				<u> </u>	IQGTCYRGKAKCCK
QCGIWVRTSNGGHFASPNYPDSYPPNKECIYI LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMKGSRFKAFSTREDAEKFAR GICDYFPSSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRRYVVDLYLNTP DKMGYDTPLHFACKFGNADVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHYSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	888	2238	A	7702	242	1298	
LEAAPRQRIELTFDEHYYIEPSFECRFDHLEVR DGPFGFSPLIBTYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEBAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVVEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	1		1			}	
DGPFGFSPLIDRYCGVKSPPLIRSTGRFMWIKF SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEBAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFA GICDYPPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEFYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	1		1	·			, ,
SSDEELEGLGFRAKYSFIPDPDFTYLGGILNPIP DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAGGNFTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGGREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA			1		1		
DCQFELSGADGIVRSSQVEQEEKTKPGQAVD CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSPCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEBAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA			1				
CIWTIKATPKAKIYLRFLDYQMEHSNECKRNF VAVYDGSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEBAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	Í		(1		1	
VAVYDGSSSIENLKAKFCSTVANDVMLKTGI GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEBAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA				1			
GVIRMWADEGSRLNRFRMLFTSFGGASPAQA ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEBAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYQVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSFVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA		Ì			ŀ		
ALSFCHSNMCINNSLVCNGVQNCAYPWDEN HC 889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	1	1	ļ	1	1		
B89 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA			1				
889 2239 A 7707 185 2911 CHYIMNPSTHHPASAGGSILGLFDFFGLGLGE MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA					!	1	
MTMDALLARLKLLNPDDLREEIVKAGLKCGP ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHIVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	889	2239	A	7707	185	2911	L
ITSTTRFIFEKKLAQALLEQGGRLSSFYHHEA GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	"		1	''''			
GVTALSQDPQRILKPAEGNPTDQAGFSEDRDF GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPJFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHIVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	}	}	l	1	1	1	
GYSVGLNPPEEEAVTSKTCSVPPSDTDTYRAG ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLITLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHIVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA			1		l		
ATASKEPPLYYGVCPVYEDVPARNERIYVYE NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA		1	1	1	1		
NKKEALQAVKMIKGSRFKAFSTREDAEKFAR GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA]		J.			j	
GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG LCLSESETVNKERANSYKNPRTQDLTAKLRK AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAETSSPVIGELWSPQOTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	1	1	1	1	1		
AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA]	}			1		GICDYFPSPSKTSLPLSPVKTAPLFSNDRLKDG
AVEKGEEDTFSDLIWSNPRYLIGSGDNPTIVQ EGCRYNVMHVAAKENQASICQLTLDVLENP DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	1		1		[LCLSESETVNKERANSYKNPRTQDLTAKLRK
DFMRLMYPDDDEAMLQKRIRYVVDLYLNTP DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHIVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	1	1	1	1	l	Ì	
DKMGYDTPLHFACKFGNADVVNVLSSHHLI VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	1		Ī	,	1	[EGCRYNVMHVAAKENQASICQLTLDVLENP
VKNSRNKYDKTPEDVICERSKNKSVELKERIR EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	1	1	1		1		
EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA			1	1	1	1	DKMGYDTPLHFACKFGNADVVNVLSSHHLI
EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA			1		[1	
EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE DFRKLWKTPPREKAGFLHHVKKSDPERGFER VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA		1	1	1	1	1	EYLKGHYYVPLLRAEETSSPVIGELWSPDQTA
VGRELAHELGYPWVEYWEFLGCFVDLSSQE GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	}		l .	1			EASHVSRYGGSPRDPVLTLRAFAGPLSPAKAE
GLQRLEEYLTQQEIGKKAQQETGEREASCRD KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	İ	ł	1	1	l	i	
KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	1		ļ		1		VGRELAHELGYPWVEYWEFLGCFVDLSSQE
KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA	1	1	1	1		1	
ARNNSPPTVGAFGHTRCSAFPLEQEADLIEAA	1	1		1	1	1	KATTSGSNSISVRAFLDEDDMSLEEIKNRQNA
	L	<u> </u>	<u></u>		<u> </u>	<u></u>	ARNNSPPTVGAFGHTRCSAFPLEQEADLIEAA

891 2241 A 7721 61 1175	SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EPGGPHSSRNGLCHPLNHSRTLAGKRPKAPR GEEAHLPPVSDLTVEFDKLNLQNIGRSVSKTP DESTKTKDQILTSRINAVERDLLEPSPAQLG NGHRRTESEMSARIAKMSLSPSSPRHEDQLEV TREPARRLFLFGEEPSKLDQDVLAALECADV DPHQFPAVHRWKSAVLCYSPSDRQSWPSPAV KGRFKSQLPDLSGPHSYSPGRNSVAGSNPAKP GLGSPGRYSPVHGSQLRRMARLAELAAL
SYSQYEKLDAGEGRIMDEAFORDEGETIL HSPSDWITSHPEAPQDEGETSPYRKTPSPN KRSIYIQSIGSLGNTRIISEEVIK WLTGYCK AYF YGLRVKLLEPVPVSVTRCSFRVNENTHNLQIH AGDILKFLKKKPEDAFCVVGITMIDLYPRDS WNFVFQASLTDGVGIFSFARYGSDFYSMFY KGKVKKLKKTSSDYSIFDNYIPPEITSVILLR SCKITTHEIGHIFGLRHCQWLACLMGGSNHL EEADRRIPLN.CPICHKLQCAFSIVERYKA LVRWIDDESSDTFGATPEHSHEDNGNLPKPV EAFKEWKEWIKCLAVLQK A 7723 2 1650 SAPTAPARPCRAERGSGGGMLALLAASVALA VAAGAQDSPAFGSRFVCTALPFEAVHAGCFL PAMPMQGGAQSFEELRAAVLQL PAMPMQGGAQSFEELRAAVLQL RETAYQQKETLASARAIRELTGKLARCEGLAGGKARGA GATGKDTMGDLPROPGHYVEQLKETVVQQ KETLASARAIRELTGKLARCEGLAGGKARGA GATGKDTMGDLPROPGHYVEQLKETVVQQ KETLASARAIRELTGKLARCEGLAGGKARGA GATGKDTMGDLPROPGHYVEQLKETVYQQ KETLASARAIRELTGKLARCEGLAGGKARGA GATGKDTMGDLPROPGHYVEQLKESTLYQL LRETVYQQKETLASARAIRELTGKLARCEGLAGGKARGA GATGKDTMGDLPROPGHYVEQLKESTLYQL LSRSLGTLKDRLESLEHQLRANVSNAGLPGD FREWLQQRLGGLERQLLRKGAELEDEKSLLH NETSAHRQKTESTLNALLQRYTELERGINSAF KSPNAFKVSLPLRTNYLYGKGAELEDEKSLLH NETSAHRQKTESTLNALLQRYTELERGINSAF KSPNAFKVSLPLRTNYLYGKGAELEDEKSLLH NETSAHRQKTESTLNALLQRYTELERGINSAF KSPNAFKVSLPLRTNYLYGKGAELEDEKSLLH RYGGVLLGQEQDTVGGRFDATQAFVGLSQ FINWBVLRAQQIVNIANCSLGTGENLAWHITO TITRDGMWAAPQDKKLGTGENLAWHITO KYGGVLLGQEQDTVGGRFDATQAFVGLSQ FINWBVLRAQQIVNIANCSLGTGENLAWHITO KYGGVLLGQEQDTVGGRFDATQAFVGLSQ FINWBVLRAQQIVNIANCSLGTGENLAWHITO KYGGVLLGQEQDTVGGRFDATQAFVGLSQ FINWBVLRAQQIVNIANCSLGTGENLAWHITO LCKTVIALIKVNFYCSSLLLACLAVBSGWVLGTF LCKTVIALIKVNFYCSSLLLACLAVBSGWVLGTF LCKTVIALIKVNFYCSSLLLACLAVBYGWTHAWF TSRFLYHVAGFILPMLVVGFLALPL LFAKVSQGHINNSLPRCTFSQENQAETHAWF TSRFLYHVAGFILPMLVVGFLALPL QAQRRPQRQXAVRVALLVTSIFFLCWSPYHIV GLAICCLNPMLYTFACVKRSDLSRLITKLG CLAICCLNPMLYTFACVKRSDLSRLITKLG CLAICLANCHMUTYTFACVKRSDLSRLITKLG CLAICCLNPMLYTFACVKRSDLSRLITKLG CLAICCLNPMLYTFACVKRSDLSRLITKLG CLAICCLNPMLYTFACVKRSDLSRLITKLG CLAICCLNPMLYTFACVKRSDLSRLITKLG CTGPASLCQLFFSWRSSLSSESENTSLITTF CTFRAGRWGAGARVRGGAGARVGGGARXGGGR SLYPSRSVIVTRSGALRFVFKNSDLSRLITKLG CLAICCLNPMLAGAARVRGGARVGGGR SLYPSRSVIVTRSGALRFVF	890	2240	Α	7711	360	269	RHMPVIPALWEAEVGGLLEPRSSRSAWATE
VAAGAQDSPAPGSRFVCTALPPEAVHAGCPL PAMPMQGGAQSPEEELRAAVU,QLRETTVVQQ KETLASARAIRBLTGKLARCEGLAGGKARGA GATGKDTMGDLPRDPGHVVEQLSRSLQTLK DRLESLEFLFAMPMQGGAQSPEEELRAAVLQ LRETVVQQKETLASARAIRELTGKLARCEGL AGGKARGAGATGKDTMGDLPRDPGHVVEQ LSRSLQTLKDRLESLEHQLRANVSNAGLPGD FREVLQQRLGELERQLLRKGAELEDEKSLLH NETSAHRQKTESTLNALLQRVTELERGNSAF KSPNAFKVSLPLRTNYLYGKIKKTLPELVAFT ICLWLRSSASPGMGTPFSYAVPGQANEIVLIE WGNNPIELLINDKVAQLPLFVSDGKWHHICV TWTTRDGMWEAFQDGKKLGTGENLAPWHPI KPGGVLILGQEQDTVGGRFDATQAFVGELSQ FNIWDRVLRAQEIVNIANCSTNMPGNIIPWVD NNVDVFGGASKWPVETCEERLLDL DNYNDTSLVENHLCPATEGFLMASFKAVFVP VAYSLIFLLGVIGNVLVLVILERHRQTRSSTET FLFHLAVADLLLVFLPFAVAEGSVGWVLGTF LCKTVIALHKVNFYCSSLLLACLAVDRYLAIV HAVHAYRHRRLLSIHITCGTIWLVGFLLALPEI LFAKVSQGHHNNSLPRCTFSQENQAETHAWF TSRFLYFIVAGFLLPMLVMGWCVYGVVIRFLR QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV IFLDTLARIKAVDNTCKLNGSLPVATTMCEFL GLAIICCLNPMLYTFAGVKFFRSDLSRLLTKLG CTGPASLCQLFPSWRRSSLSESENATSLTTT 894 2244 A 7738 670 287 FVTRAGRWGAGARVRGGAGGMASGARWL VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR SLVPSRSVIVTRSGALLPKVKMSFGGLLRVFSI VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD D 895 2245 A 7753 119 278 APYAHSQVHCLDKVCGLLPFLNPEVPDQFYYR	891	2241		//21	61	1175	VSQYEKLDAGEQRLMNEAFQPASDLFGPITL HSPSDWITSHPEAPQDFEQFFSDPYRKTPSPN KRSIYIQSIGSLGNTRIISEEYIKWLTGYCKAYF YGLRVKLLEPVPVSVTRCSFRVNENTHNLQIH AGDILKFLKKKKPEDAFCVVGITMIDLYPRDS WNFVFGQASLTDGVGIFSFARYGSDFYSMHY KGKVKKLKKTSSSDYSIFDNYYIPEITSVLLLR SCKTLTHEIGHIFGLRHCQWLACLMQGSNHL EEADRRPLNLCPICLHKLQCAVGFSIVERYKA LVRWIDDESSDTPGATPEHSHEDNGNLPKPV
DNYNDTSLVENHLCPATEGPLMASFKAVFVP VAYSLIFLLGVIGNVLVLVILERHRQTRSSTET FLFHLAVADLLLVFILPFAVAEGSVGWVLGTF LCKTVIALHKVNFYCSSLLLACIAVDRYLAIV HAVHAYRHRRLLSIHITCGTIWLVGFLLALPEI LFAKVSQGHHNNSLPRCTFSQENQAETHAWF TSRFLYHVAGFLLPMLVMGWCYVGVVHRLR QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV IFLDTLARLKAVDNTCKLNGSLPVAITMCEFL GLAIICCLNPMLYTFAGVKFRSDLSRLLTKLG CTGPASLCQLFPSWRRSSLSESENATSLTTF 894 2244 A 7738 670 287 FVTRAGRWGAGARVRGGAGGMASGAARWL VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD D 895 2245 A 7753 119 278 APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR						,	SAPTAPARPCRAERGSGGGMLALLAASVALA VAAGAQDSPAPGSRFVCTALPPEAVHAGCPL PAMPMQGGAQSPEEELRAAVLQLRETVVQQ KETLASARAIRELTGKLARCEGLAGGKARGA GATGKDTMGDLPRDPGHVVEQLSRSLQTLK DRLESLEPLPAMPMQGGAQSPEEELRAAVLQ LRETVVQQKETLASARAIRELTGKLARCEGL AGGKARGAGATGKDTMGDLPRDPGHVVEQ LSRSLQTLKDRLESLEHQLRANVSNAGLPGD FREVLQQRLGELERQLLRKGAELEDEKSLLH NETSAHRQKTESTLNALLQRVTELERGNSAF KSPNAFKVSLPLRTNYLYGKIKKTLPELYAFT ICLWLRSSASPGMGTPFSYAVPGQANEIVLIE WGNNPIELLINDKVAQLPLFVSDGKWHHICV TWTTRDGMWEAFQDGKKLGTGENLAPWHPI KPGGVLILGQEQDTVGGRFDATQAFVGELSQ FNIWDRVLRAQEIVNIANCSTNMPGNIIPWVD
894 2244 A 7738 670 287 FVTRAGRWGAGARVRGGAGGMASGAARWL VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD D 895 2245 A 7753 119 278 APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR	893	2243	A	7729	3554	2419	LTAGTAMNYPLTLEMDLENLEDLFWELDRL DNYNDTSLVENHLCPATEGPLMASFKAVFVP VAYSLIFLLGVIGNVLVLVILERHRQTRSSTET FLFHLAVADLLLVFILPFAVAEGSVGWVLGTF LCKTVIALHKVNFYCSSLLLACIAVDRYLAIV HAVHAYRHRRLLSIHITCGTIWLVGFLLALPEI LFAKVSQGHHNNSLPRCTFSQENQAETHAWF TSRFLYHVAGFLLPMLVMGWCYVGVVHRLR QAQRRPQRQKAVRVAILVTSIFFLCWSPYHIV IFLDTLARLKAVDNTCKLNGSLPVAITMCEFL GLAIICCLNPMLYTFAGVKFRSDLSRLLTKLG
895 2245 A 7753 119 278 APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR	894	2244	A	7738	670	287	FVTRAGRWGAGARVRGGAGGMASGAARWL VLAPVRSGALRSGPSLRKDGDVSAAWSGSGR SLVPSRSVIVTRSGAILPKPVKMSFGLLRVFSI VIPFLYVGTLISKNFAALLEEHDIFVPEDDDDD
	895	2245	Α	7753	119	278	APYAHSQVHCLDKVCGLLPFLNPEVPDQFYR
	896	2246		7754		372	LWLSLFLHAGKEAPHCPRTRPL SPAWWNSQQRVVSPFLALLTLEPTFHHLLPIM

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion QVSTAALAVLLCTMALCNQVLSAPLAADTPT ACCFSYTSRQIPQNFIADYFETSSQCSKPSVIFL
897	2247	A	7761	1725	445	TKRGRQVCADPSEEWVQKYVSDLELSA RPRRRGTHHFSCVLGSFRVSAMFPRVSTFLPL RPLSRHPLSSGSPETSAAAIMLLTVRHGTVRY RSSALLARTKNNIQRYFGTNSVICSKKDKQSV RTEETSKETSESQDSEKENTKKDLLGIIKGMK VELSTVNVRTTKPPKRRPLKSLEATLGRLRRA TEYAPKKRIEPLSPELVAAASAVADSLPFDKQ TTKSELLSQLQQHEEESRAQRDAKRPKISFSNI ISDMKVARSATARVRSRPELRIQFDEGYDNYP GQEKTDDLKKRKNIFTGKRLNIFDMMAVTKE APETDTSPSLWDVEFAKQLATVNEQPLQNGF EELIQWTKEGKLWEFPINNEAGFDDDGSEFH EHIFLEKHLESFPKQGPIRHFMELVTCGLSKNP YLSVKQKVEHIEWFRNYFNEKKDILKESNIQF KLRPWKFLFRNN
898	2248	Α	7775 •	85	496	SCOTTOPPAQSCSTGTMRIMLLFTAILAFSLA QSFGAVCKEPQEEVVPGGGRSKRDPDLYQLL QRLFKSHSSLEGLLKALSQASTDPKESTSPEK RDMHDFFVGLMGKRSVQPDSPTDVNQENVP SFGILKYPPRAE
899	2249	A	7785	179	703	PFHLGASSNTFRLQVQTQESKAQKEVKMGFI FSKSMNESMKNQKEFMLMNARLQLERQLIM QSEMRERQMAMQIAWSREFLKYFGTFFGLA AISLTAGAIKKKKPAFLVPIVPLSFILTYQYDL GYGTLLERMKGEAEDILETEKSKLQLPRGMIT FESIEKARKEQSRFFIDK
900	2250	A	7789	1465	300	VWLPLKSYKIRSPSLHCQCEIFREEFLFSSLQE GRDKDTFSKMAMVSEFLKQAWFIENEEQEY VQTVKSSKGGPGSAVSPYPTFNPSSDVAALH KAIMVKGVDEATIIDILTKRNNAQRQQIKAAY LQETGKPLDETLKKALTGHLEEVVLALLKTP AQFDADELRAAMKGLGTDEDTLIEILASRTN KEIRDINRVYREELKRDLAKDITSDTSGDFRN ALLSLAKGDRSEDFGVNEDLADSDARALYEA GERRKGTDVNVFNTILTTRSYPQLRRVFQKY TKYSKHDMNKVLDLELKGDIEKCLTAIVKCA TSKPAFFAEKLHQAMKGVGTRHKALIRIMVS RSEIDMNDIKAFYQKMYGISLCQAILDETKGD YEKILVALCGGN
901	2251	A	7796	2	807	VEFHPQRARAGARAPSMGVLLTQRTLLSLVL ALLFPSMASMAAIGSCSKEYRVLLGQLQKQT DLMQDTSRLLDPYIRIQGLDVPKLREHCRERP GAFPSEETLRGLGRRCFLQTLNATLGCVLHRL ADLEQRLPKAQDLERSGLNIEDLEKLQMARP NILGLRNNIYCMAQLLDNSDTAEPTKAGRGA SQPPTPTPASDAFQRKLEGCRFLHGYHRFMH SVGRVFSKWGESPNRSRRHSPHQALRKGVRR TRPSRKGKRLMTRGQLPR
902	2252	A	7802	2	721	TAARRRQKGTAARRLQKGTAARRRQKGTAA RRRQKGTAARRPQKGTAARRRQKGTAARRR QKGTAARRRQKGTAARRPQKGTAARRRQKG TAARRRQKGTAARRRQKGLAIASRGCPCASR AGGVRGAGSRLRAMAPKVFRQYWDIPDGTD CHRKAYSTTSIASVAGLTAAAYRVTLNPPGTF LEGVAKVGQYTFTAAAVGAVFGLTTCISAHV REKPDDPLNYFLGGCAGGLTLGARTHNYGIG AAACVYFGIAASLVKMGRLEGWEVFAKPKV

SOUT Doc	Cero in	SEQ ID	T Man	SEQ	Thus 37 - 4 - 3	130	
nuclocide seq uence unce unce unce unce unce unce unce	SEQ ID		Met		Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence USAN O9496 corresponding to last amino acid residue amino acid residue of peptide residue of peptide sequence peptide s			nou .	1			
Seq		1			5 .		
1914 18 antion acid residue of peptide peptide peptide peptide peptide peptide peptide sequence 1917 1918			}			to last amino	M=Methionine N=A charagine P=Proline
mino aed reidue of peptide requence purpose sequence purpose sequence purpose sequence purpose sequence purpose sequence provide sequence provide sequence provide sequence provided sequence sequence provided sequence provided sequence provided sequence sequence provided sequence pr	1 -		1	1			
Pesidue of pepide Pepide		1					T=Threonine, V=Valine, W=Tryptonhan
Pepilde		ŀ					Y=Tyrosine, X=Unknown, *=Stop codon.
903			ļ	}	peptide		
903	l	ł	ł	l		1	
904 2254 A 7813 40 821 GAGRALGHETGAGDVAAALPARKFPESLLO LIFEKPR REDST/DSQKDMEIPJPWQRTDESISTR ARLLYESRKRGMLENCILLSLFAKEHLQHM EQUIL YDRILINPESNDWDYVWATAEAPAP EIFENEVMALLRIPAKNKNKEQRLRAPDLEY LIFEKPR LIFE	903	2253	Α	7807	1	584	
RGDSPIDSQKDMEIPLPPWQERTDESISTER ARLLYSERRGMLEROFLLISLARAEHQHM EKQLMLYDELINSPERROMDIPYWATEAKPAP EIFENEVMALLRPAKDWDIYWATEAKPAP EIFENEVMALLRPAKDWDIYWATEAKPAP LIFEKPR 904 2254 A 7813 40 821 GAGRALGHLETGAGDVAAALPARKFPSILG AGARLGWIMNVFRILGDLSHLLAMILLIGK IWRSKCCKGISGKSQILFALVFITRYLDLFTNF BISTNTVMKVYFLLGALYTTRYLDLFTNF DSENDIFFRLIFFLLVFVIGLSFLENYSFTLLEIL WTFSITLSSVALRQLFMSKTGEASTITHYL FFLGLYRALYLANWIRRYQTENPYDQIAVVS GVVQITFSCDFTLVFVTGGSSPLENYSFTLLEIL WTFSITLSSVALRQLFMSKTGEASTITHYL FFLGLYRALYLANWIRRYQTENPYDQIAVVS GVVQITFSCDFTAMSTGEASTITHYL FFLGLYRALYLANWIRRYQTENPYDQIAVVS GVVQITFSCDFTAMSTGEASTITHYL FFLGLYRALYLANWIRRYQTENPYDQIAVVS QVVQINEVDEIMSKTGASTITHYL FFLGLYRALYLANWIRRYQTENPYDQIAVVS QVVQINEVDEIMSKTGASTITHYL FFLGLYRALYLANWIRRYQTENPYDQIAVVS QVVQINELDIAMSTROGGASLTVVKQNAD VALONIRVMNSAQASIEQLVSGASTILNIVA EILKSIDRISEVKDESEDS 906 2256 A 7822 3 1462 DSPRINFELIGRPTRIPTRPGFPRAMEDLDAL LISDLETITSIMPRSGAFKERPAEPLTPPSVG HQPQTISGESSAGSGDKDHLYSTVCKPRSPK PAAPAAPPSSSSGVLOTGLGELDRALIQELNA TQPNITDEIMSQFPSSKVASGGKEPQSEDKK RPSLSSSFGGFRASTSATLELDRIMASLSD FRVQNIHLPASGPTQPVVSSTSTSGSPSPPETTG KGSLDTMCGLUGSDLSRRGVPTQAKCLGSS NKPIAGQVVTALGRAWHPEHPVCGGCSTAL GGSSFFEEGDAGPFCECVFFERSPRGCPCNOPI RHKMYTALGTHWHPEHPCCVSGCSTAL GGSSFFEEGDAGPFCEVCFFERSPRGCPCNOPI RHKMYTALGTHWHPEHPCCVSGCSTAL GGSSFFEEGDAGPFCEVCFFERSPRGCPCNOPI RHKMYTALGTHWHPEHPCCVSGCSTAL GGSSFFEEGDAGPFCEVCFFERSPRGCPCNOPI HTVNQSFASPSDQEVGTLYECFGSDGKLVLH YCKSQAWG 2258 A 7842 110 11172 KLSCSGGRTSHEE GRPLENHFHARGLSCLATGLEVTGRCVSA LGRRFHPDHFTCFCLRPLTKGSFQERAGKPY CQVCFLKLFG TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CLNATEWREHASGVYSRSDTSSSTTGSTVQ SVDLFFTRLWTYSHHTYNKCKRKNILEWAKEL SLGFSSWGFGFGVGVCVGPGSACEEFWARLR KLNWKRILRIREDPFDGTIDGTERQKRSFIG TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV SVDLFTRLWTYSHHTYNKCKRKNILEWAKEL SLGFSFWGFGFGVCVCGPQSACEEFWARLR KLNWKRILRIREDPFDGTDGTDGTERQKRSFIG TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV SVDLFTRLWTSHHTYNKCKRKNILEWAKEL SLGFSFWGFGVGVCVCGPQSACEEFWARLR KLNWKRILRIREDPFDGTDGTDGTBCQRTSTGSTQTO SVDLFTRLWTSHTTYNCCRKKNILEWAKEL SLGFSFWGFGVGVCVCGPQSACEEFWARLR KLNWKRILRIREDPFDGTDGTDGTERQKFSFIG TVRSVLLSRSQQTQLTDFT	1						VSTVFSTSSLMLALSRHSLLSPLLSVTSFRRFY
ARLLYBERKERMLENCILLSLFAKEHI,QHMT ERQLIN_YDRLINIPSROWDIYYWARFAKPA EIEFENEYMALLRDFAKNKNKEQRLRAPDLEY LFEKPR LFEKP	1		1			ļ	
		i	-		1	1	
BEFENEYMALLRDFAKNINEQRLRAPDLEY FERPR			İ		1		EKQLNLYDRLINEPSNDWDIYYWATEAKPAP
1		ļ	}]	į	EIFENEVMALLRDFAKNKNKEQRLRAPDLEY
AGARLTGWTMNVFRLIGDLSHILAMILLIGK IWRSKCKGISGKSQILFALAWFITRYLDLFTNT SIYNTVMEVVFLICAYTVTYMIVGKFRKIT DSENDIFRLEFLLVPVIGLSFLENYSFILLEIL WITSYLESVALIPQ-FMSKTGLEAFITITHYL FFLGLYRALYLANWIRRYQTENFYDQIAVVS GVVQIIFYCDFFYLYVKGSSWDDSNADTGL RSYSSI SINKDVLSPQLKDENSKLRKLNEVQSFSEA QTEMVRTLERKLEAKMIKEESDYHDLESVVQ QVEQNLEMTKRAVKAARNVVKLKQIEILL QAQVSNPQRENEARCQGGASILTVVKQNAD VALQNLRVVMNSAQASIBQLVSGAETLNLVA ELKSENSISVENDEEDS DSPRNRFEILGRFTRTFTRFORPAMEDIDAL SINKDVLSPQLKDEEDS DSPRNRFEILGRFTRTFTRFORPAMEDIDAL SINKDVLSPQLKTSTLENDA TOPNITDELBAGPFSSKVASGEQKEDQSEDKK RPSLPSSPSPGLKGASATSATLELDRLMASLSD FRVQNHLPASGFTQFVVSSTNEAGSPFPETTG KGSLDTMLGILQSDLSRRGYTQAKGGCGSTAL GGSSFEKDGAPFCPECYFERSPRGGFCOPP RHKMYTALGTHHPEHFCVSCGEFFGDEG FHERGORPYCRRDFLQLFARQGCQGFLDU SINKDVALGVETALGERSPPETTG KGSLDTMLGILQSDLSRRGYTQAKGCCGFFGDEG FHERGORPYCRRDFLQLFARQGCQGFLDU SINKDVALGVETALGERSPPETG KGSLDTMLGILQSDLSRRGYTQAKGCCGFFGDEG FHERGORPYCRRDFLQLFARQGCQGFLDU SINKDVALGVETALGERSPPETG KGSLDTMLGILQSDLSRRGYTQAKGCCGFFGDEG FHERGORPYCRRDFLQLFARQGCQGFLDU SINKDVALGVETALGERSPPETG KGSLDTMLGILQSDLSRRGYTQAKGCCGFFGDEG FHERGORPYCRRDFLQLFARQGCQGFLDU SINKDVALGVETALGERSPPETG KGSLDTMLGILGSDLSRRGYFTQAKGCCGFTGDEG FHERGORPYCRRDFLQLFARQGCQGFLDU SINKDVALGVETALGEFTGATG SINKDVAL	<u></u>					<u> </u>	
WRSKCCKGIGGKSQULFALVTTRYLLDLTNF SIYNTYMKVVPLLCAYYTVYMIYGKERKTF DSENDIFRLEFLLVPVIGLSFLENYSFILLEIL WTFSTYLESVALLPQLFMISKTGEATITITHYL FFLGLYRALVLANWIRKYQTENFYDQIAVVS GVVQITPYCDFFYLLYTKGRSWDDSNADTGL RSYSSI SINKDVLSPQLKDENSKLRRKINEVQSFSEA QRWRITERLEAKMIKESBYDHDLESVVQ QVEQNLEIMTKRAVKAENHVVKLKQEISLL QAQVSNFQRENEALRCQGASLTVVKQNAD VALQNLRVVMNSAQASIEQLVSGAETLNLVA ELKSIDRISEVKDEEEDS DSPRNFFELIGRFTIFTRPGFRPAMEDLDAL LSDLETTTSHMPRSGAFKERPAEPTPPSYG HQPQTGSGESSGASGDKDHLYSTVCKPRSPK PAPA-PRASPFSSSGVLGTGLCELDRLLQELNA TQFNITDEIMSQFPSSK VASGEQKEDQSBKK RRSLPSSPSFOLPKASSTSATLEDRLMASLSD FRVQNHLPASGPTQPFVVSSTNEGSSSPEPTG KGSLDTMLGULGSLSRRGVPTQAKGLCGSC NKPIAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFKKDGAPFCPEVTERSPRCGFCNQPI GRIKMYTALGTHWHPEHFCVSCGEPFGDEG FHERRGRPVCRFFLQLFLARGCVSA LGRRRHDHFTCTFCLRPLTKGSFGERAGKPY CQPCFLKLFG PIYNOSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG VYCKSQAWG VLSSGARGHADGGQLVGTLD VTCKMEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCKKTMEGSKVYFTI NNMLDVSDEKMAMFSLACILPFKYPAVLPBI TVRSVLLSRSQTQLNTDLTAFLQKRCHGDV CLNATEWVREHASGVVSRDTSSPTTGSTVQ SVDLIFTRL WYSHHYNCKRSKNLEWAKLR KLNWKRILRIRREDIFFDGTNDETERQRKFSIF EKVFSVNGARGNHMDFGGLYCHTKCGC DVFQMFLWV SUDLIFTRL WYSHHYNCKRKNILEWAKLR KLNWKRILRIRREDIFFDGTNDETERQRKFSIF EKVFSVNGARGNHMDFGGLYCHTKCGC DVFQMFLWV SUDLIFTRL WYSHHYNCKRKNILEWAKLR KLNWKRILRIRREDIFFDGTNDETERQRKFSIF EKVFSVNGARGNHMDFGGLYCHTKGCO DVFQMFLWV SUDLIFTRL WYSHHYNCKRGCHCKMEDWAKLR KLNWKRILRIRREDIFFDGTNDETERQRKFSIF EKVFSVNGARGNHMDFGGLYCHTKTCGC DVFQMFLWV SUDLIFTRL WYSHHYNCKRGCHCHGDV CRASCH SURLIFRC SURL	904	2254	A	7813	40	821	GAGRALGHLETGAGDVAAALPARKFPRSLLG
SISYNTVMKVVFLICAYVTVYMINGKFRKTF DESENDIFRLEFILVPVIGLS ENYSFTILEII. WTFSIYLESVALIPQLEMISKTGEAETITITHVL FFLGLYRALYLANWIRKYQTENYSTILEII. WTFSIYLESVALIPQLEMISKTGEAETITITHVL FFLGLYRALYLANWIRKYQTENYDIAVVS GVVQITFYCDFFYLVYTKGRSWDDSNADTGL RSYSSI	ļ		ŀ				
DSENDITFILEFILLIPYVIGLSFILENSYFILLEIL WTPSIYLESVALIPQLBMSKTGAETITTHYLL FFLGLYRALYLANWIRRYQTENFYDQIAVVS GVVQIITYCDFFYLYVTKGRSWDDSNADTGL RSYSSI			!	•			IWRSKCCKGISGKSQILFALVFTTRYLDLFTNF
WTFSIYLESVALLPQLEMISKTGEAETITITH/L FELGLYRALVLANWERVGPTPQLIAVVS GVVQITFYCDFPYLYVTKGRSWDDSNADTGL RSYSSI		1				ļ	ISIYNTVMKVVFLLCAYVTVYMIYGKFRKTF
PFLGLYRALYLANWIRRYQTENFYDQIAVVS GVVQTIFYCDFYLYVTKGRSWDSNADTGL		i					1
905 2255 A 7817 1399 881 LSNKDVLSPQLKDENSKLRRKLNEVQSFSEA QTEMVRTLERKLEAKMIKESDYPHDLESVVQ QVEQNLELMTKRAVKABNIVVKIKQEISLL QAQVSNFQRENRALRCGQGASLTVVKQNAD VALQNLRVVMNSAQASIEQLVSGAFLINLVA EILKSIDRISEVKDEEEDS 2256 A 7822 3 1462 DSFRINFEILGRFTRTPTRFORPRAMEDLDAL EILKSIDRISEVKDEEEDS DFRINFEILGRFTRTPTRFORPRAMEDLDAL TOPINTDEIMSQFPSSKVASEQKEDQSEDKK RPSLPSSPSFGLPKASATSATLELDRLMASLSD FRVQNILPASGFTQPPVVSSTNEGSPSPPPTG KGSLDTMLGLLQELNA TOPINTDEIMSQFPSSKVASEQKEDQSEDKK RPSLPSSPSFGLPKASATSATLELDRLMASLSD FRVQNILPASGFTQPPVVSSTNEGSPSPPPTG KGSLDTMLGLLQELNA GGSSFFKLDGAPFCPPCYPERFSPRCGFCNQPI RIKMVTALGTHWHPEHFCCVSCGESTAL GGSSFFKLDGAPFCPPCYPERFSPRCGFCNQPI RIKMVTALGTHWHPEHFCCVSCGESTAL GGRSFFKLDGAPFCPPCYPERFSPRCGFCNQPI RIKMVTALGTHWHPEHFCVSCGESTAL GGRSFFKLDGAPFCPPCYPERFSPRCGFCNQPI RIKMVTALGTHWHPEHFCVSCGESTAL GGRSFFKLDGAPFCPPCYPERFSPRCGFCNQPI RIKMVTALGTHWHPEHFCVSCGESTAL GGRSFFKLDGAPFCPPCYPERFSPRCGFCNQPI RIKMVTALGTHWHPEHFCVSCGESTAL GGRSFFKLDGAPFCPPCYPERFSPRCGFCNQPI RIKMVTALGTHWHPEHFCVSCGESTAL GGRSFFKLDGAPFCPPCYPERFSPRCGFCNQPI RIKMVTALGTHWHPEHFCVSCGSTAL GGRSFFKLDGAPFCPPCYPERFSPRCGFCNQPI RIKMVTALGTHWHPEHFCVSCGSTAL GRPLCENHFHARGSLCATCGLPVTIRCVSA LGRRHPDPHFCTTCLRPLTKGSFQERAGRPY COPCTLKLFG FIYVNGSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG SLOPPSGLKGSSHLSSSSWDFRHAPTHPET YCKSQAWG SLOPPSGLKGSSHLSSSWDFRHAPTHPET YCCPKMIEMEQAEAQLABLDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACLPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKFGKYPGFLKAGVQWLDG SVDLIFTRLWYSHHYNKCKRKNILEWAKEL SLSGFSMPGRGVVCVEGPQSACEEFWARLR KLNWKRLIRIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHDFGQLYQFLNTKGCG DVPQMFLWV GRATWHPTATCHTYPYMYMNSV LISSELLIPSKYLFESK GALTWSHPLLAVCPQQVWLGSTPSGSPALLP PSHRVNABPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPQYYGGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVQGYFGHHCCHRMD SCRHLPGAPHGYTCDCVQGYFGHHCCHRMD SCRHLPGAPHGYTCDCVQGYFGHHCCHRMD SCRHLPGAPHGYTCDCVQGYFGHHCCHRMD SCRHLPGAPHGYTCDCVQGYFGHHCCHRMD SCRHLPGAPHGYTCDCVQGYFGHHCCHRMD SCRHLPGAPHGYTCDCVQGYFGHHCCHRMD SCRHLPGAPHGYTCDCVQGYFGHHCCHRM		ļ					
905 2255 A 7817 1399 881 LSNKDVLSPQLKDENSKLRRKLNEVQSFSEA QTEMVRTLERKLEAKMIKEESDYHDLESVVQ QVEQNLELMTKRAVKAENHVVKLKQEISLL QAQVSNFQRENEALRGQGASLTVVKQNAD VALQNLRVVMNSAQASIEQLVSGAETLNLVA ELLKSIDRISEVKDEEEDS			ļ .				FFLGLYRALYLANWIRRYQTENFYDQIAVVS
905 2255 A 7817 1399 881 LSNKDVLSPQLKDENSKLRKKLNEVQSFSEA QTEMVRTLERKLAKMIKEESDYHDLESVVQ VEQNLELMTKRAVKAENHVVKLKQEISILL QAQVSNFQRENEALRCQGASLTVKKQNAD VALQNLRVMNSQAASIQLVSGASTLNLVA EILKSIDRISEVKDEEEDS 906 2256 A 7822 3 1462 DSFRNRFEILGRFTRTFTRFGRPAMEDLDAL LSDLETITISHMPRSGAPKERPAEPLTPPSVG HQPQTGSGESSGASGOKDHL.YSTVCKPRSPK PAAPAAPPFSSSSGVLGTGLECLDRLLQELNA TQPNITDEIMSQFPSSKVASGEKEDQSEDKK RPSLPSSPSGULFKASATSATLELDRLMASLSD FRVONHLPASGPTQPVVSSTNEGSPSPPPTIG KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPIAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECVFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFCVSGEPFGDEG FHEREGRYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPFSGGSFFEHE GRPLCENHHHARRGSLCATCGLPVTGRCVSA LGREFHEPDHFTCTFCLTRLTKGSFQERAGRYY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIVYNGSPAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SUQPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAFAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWRIFLASGVSNDSSPTTGSTVQ SVDLIFTRLWIYSHHYNKCKRKNILEWAKEL SLSGSMPGKVVCVEGPGSACEFWARLR KLNWKRLIRRIREDIPTDGTNDETERQRKSIF EEKVFSVNGARGHIMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSELLIPSKYLFESK 910 2260 A 7884 212 4874 GALTUSKYLLPSKYLPSK							
906 2256 A 7822 3 1462 DSPRNRFEILGRPTRIPTREGRAMEDLDAL LSDLETTSHMYRSQASHEVLYKORAD VALQNLRVVMNSAQASIEQLVSGAETLNLVA EILKSIDRISEVKDEEEDS 906 2256 A 7822 3 1462 DSPRNRFEILGRPTRIPTREGRAMEDLDAL LSDLETTSHMYRSGAPKERPAEPLTPPSYG HQPOTGSGESSGASGDKDHLYSTVCKPRSPK PAAPAAPPFSSSSOVLGTGICELDRALQELNA TQFNITDELMSQFPSSKVASGEQKEDQSEDKK RPSAEPSSPGIPKASATSATLELDRLMASLSD FRVQNHLPASGPTOPPVVSSTNEGSPSPPEPTG KGSLDTMLGLLQSDLSRRGVPTJQKGLGCSG NKPJAGQVVTALGRA WHEEHFVCGGGSTAL GGSSFFEKDGAPFCPECYFERFSPRGGFCNQPI RHKMYTALGTHWHPEHFCCVSGGEPGDEG FHEREGRPYCRRDFLQLFAPRCQGCQFILDN YISALSALWHPDCFVCRECFAPFSGGSFFEHE GRPLCENHHARRGSLATGLPVTIGRCVSA LGREFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAFAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQTQTQLNTDLTAFLQKHCHGDV CULNATEWVREHAGSVSKDTSSSPTTGSIVQ SVDLIFTRLWTYSHHYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPGACEEFWARLR KLNWKRILRIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGHIMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTUSPGVPGGVVAGCASGPCPPHANCRDJ WQTFSCTCQPGYYGPGCVAGLLIPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	005	2255		7017	1200		
906 2256 A 7822 3 1462 DSPRNRFEDLORPTISTREGRENGALIVACY, AND VALQNILEWTKRAVKAENIVVKLKQEISLI, QAQVSNFQRENEALRCGGQASLTVKQNAD VALQNILRVVMNSAQASIEQLVSGAETI.NLVA EILKSIDRISEVKDEEDS 906 2256 A 7822 3 1462 DSPRNRFEDLGRYTRIFTPROFRPAMEDLDAL LSDLETTTSHMPRSGAPKERPAEPLTPPPSYG HQPQTGSGESSGASGDKDHLYSTVCKPRSPK PAAPAAPPSSSSSOVLGTGLCELDRILQEINA TQFNITDEIMSQFPSSKSVASGEQKEDQSEDKK RPAPAAPPSSSSSOVLGTGIC.CELDRILQEINA TQFNITDEIMSQFPSSKSVASGEQKEDQSEDKK RPSLPSSPSOLPKASATSATLELDRILMASLSD FRVQNHLPASGPTQPPVVSSTNEGSPSPPEPTG KGSLDTMLGILQSDLSRRGVPTQAKGLCGSC NKPIAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECVFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCROFILQLFAPRCQGCQGFILDN YISALSALWHPDICFVCRECFAPPSGGSFFEHE GRPLCENHFHARRGSLATGLPVTARCGVSA LGRRFHPDHFTCTFCLRPLTKGSFQERGKPY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIYVNGSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMEMEQAEAQLAELDLLASMFPGENS LIVNDQLAVAELKDCLEKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDDV CLINATEWVREHAGSVSKDTSSSFTIGSTVQ SVDLIFTRLWYSSHHYNKCKRNILEWAKEL SLSGFSMPGKPGVVCVEGPGACEEFWARLR KLNWKRLIRRIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGHHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSELLIPSKYLFESK 900 2250 A 7884 212 4874 GALTUSRYLLAVCPQGVWLGSTFSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLIPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCCHRMD	903	2233	A	/817	1399	881	LSNKDVLSPQLKDENSKLRRKLNEVQSFSEA
906 2256 A 7822 3 I462 DSPRNRFEILGRYTRIPTRPGRPAMEDLDAL LSDLETTSHIMPRSAKERABET LPPPSYG HQPQTGSGESSGASGDKDHLYSTVCKPRSPK PAAPAAPPFSSSSGVLGTGLGLEDGLLQELINA TQFNITDEIMSQPPSSKVASGEQKEDQSEDKK RPSLPSSSSGVLGTGLGLEDGLLQELINA TQFNITDEIMSQPPSSKVASGEQKEDQSEDKK RPSLPSSPSPGLPKASATSATLELDRLMASLSD FRVONHLPASGPTQPPVVSSTNEGSPSPPEPTG KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPIAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGFLLDN YISALSALWHPDCFCRECPAPPSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIVYNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLISRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWYSHHIYNKCKRKNILEWAKEL SLSGSFMPGKPGVCGPQSACEFWARLR KLNWKRILIRIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVPCMMF.WV 909 2259 A 7870 3067 2923 EGICVYTFIYVENYTRTCMHTYPYMYMNSV LISSELLIPSKYLFESK GALLPGAPHGYTCDCVGGYFGHHCEHRMD WYFFSCTCQPGYVGGVVACALIPCQNQG SCHLIPGAPHGYTCDCVGGYFGHHCEHRMD							QIEMVKILEKKLEAKMIKEESDYHDLESVVQ
906 2256 A 7822 3 1462 DSPRNRFEILGRYRTPTRFGFRPAMEDLDAL LSDLETTTSHMPRSGAPKERPAEPLTPPSYG HQPQTGSGESSGASGDKDHLYSTVCKPRSPK PAAPAAPPFSSSSGVLGTGLCELDRLLQELNA TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK RPSLPSSPSPGLPKASATSATLELDRLMASLSD FRVQNHLPASGPTQPPVVSSTNEGSPSPPEPTG KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPLAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHEHFCVSCGEPFGDEG FHERGRPYCRRDFLQLFAPRCQGCQFILDN YISALSALWHPDCFVCRECFAPFSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTLEPLTKGSFQERAGKPY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPFSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMEMEQAGAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILFKYFAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWYREHASGYVSRDTSSPTTGSTVQ SVDLIFTIR WIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEFWARLR KLNWKRILIRIIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSELLIPSKYLFSK GGCVPTANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGGVWLGSTPSGSPA							
906 2256 A 7822 3 1462 DSPRNRFEILGRPTRTPTRPGPRPAMEDLDAL LSDLETTTSHMPRSGAKERPAEPLTPPSYG HQPQTGSGESSGASGMDMLIYSTVCKPRSPK PAAPAAPPFSSSSG/TGGICELDRLIQELNA TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK RPSLPSSPSPGLPKASATSATLEDRLIMASLSD FRVQNHIPASGPTOPPVVSSTNEGSPPEPTG KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPIAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGGCQFCDLDN YISALSALWHPDCFVCRECFAPPSGGSFFEHE GRPLCENIFHARRGSLCATCGLPVTGRCVSA LGRRIPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMEMEQAEAQLAELDLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTGLYECFGSDGKLVLH SLSSSWDFRHAPTHPET YTCPKMEMEQAEAQLAELDLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTGLTPFKYPAVLPEI TVRSVLLSRSQQTCDLTAFFKYPAVLPEI TVRSVLLSRSQQTCDLTAFFKYPAVLPEI TVRSVLLSRSQQTCDCYCVGGQSACEFWARLR KLNWKRILIRIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVINNYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK GALFURDAFFENDE GAL							
906 2256 A 7822 3 1462 DSPRNRFEILGRPTRIPTRPGPRPAMEDLDAL LSDETTITSHMPRSGAPKEPAEPLTPPPSYG HQPQTGSGESSGASGOKDHLYSTVCKPRSPK PAPAPAPPSSSSGYLGTGLCELDRLLQELNA TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK RPSLPSSPSGPLFKASATSATLELDRLMASLSD FRVQNHLPASGPTOPPVSSTNEGSPSPPETTG KGRIDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPJAGQVVTALGRAWHPEHFVCGGCSTALL GGSSFFEKDGAPFCPECYPERFSPRCGFCNQPI RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCREDPLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPPSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRRHPDHFTCTFCLRPLTKGSFQERGKPY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQGYQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWYSHHIVNKCKRKNILEWAKEL SLSGFSMPGKRGVVCVCGFQSACEEFWARLR KLNWKRLIRIREDPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSIRRVNABPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYTGHHCEHRMD							
LSDLETTTSHMPRSGAPKERPAEPLTPPPSYG HQPQTGSGESSGASGOKDHLYSTVCKPRSPK PAAPAPPSSSSGVLGTGLCELDRLLQELNA TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK RPSLPSSPSGPLKASATSATLELDRLMASLSD FRVQNHILPASGPTQPPVVSSTNEGSPSPPEPTG KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPIAGQVVTALGRAWHPEHFVCGCGSTAL GGSSFFEKDGAPFCPGCYFGRFSPRCGFCNQPI RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPFSGGSFFEHE GRPLCENIFHARRGSLCATGGLPVTGRCVSA LGRRHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCFCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILFFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVXRSDTSSSPTITGSTVQ SVDLIFTILWTSHHTYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRLIRIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSELLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSIRRVNABPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGGCVDACLLNPCQNQG SCRRILPGAPHOYTCDCVGGYFGHHCEHRMD	906	2256	A	7822	3	1462	
HQPQTGSGESSGASGDKDHLYSTVCKPRSPK PAAPAAPPFSSSSGVLGTLELDRLLQELNA TQFNITDEIMSQFPSSKVASGEQKEDQSEDKK RPSLPSSPSPGLPKASATSATLELDRLMASLSD FRVQNIHLPASGPTQPPVVSSTNEGSSPSPEPTG KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPIAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERSPRCGFCNQPI RHKMVTALGTHWHPEHFCVCSGCSFTAL GGSSFFEKDGAPFCPECYFERSPRCGFCNQPI RHKMVTALGTHWHPEHFCVCSGCSFTAL GGSSFFEKDGAPFCPECYFERSPRCGFCNQPI RHKMVTALGTHWHPEHFCVCSGCSFTAL GGSSFFEKDGAPFCPECYFERSPRCGFCNQPI RHKMVTALGTHWHPEHFCVCSGCSFTAL GGSSFFEKDGAPFCPCCYFERSPRCGFCNQPI RHKMVTALGTHWHPEHFCVCSGCPFGDEG FHEREGRPYCRDFLQLFAPRCQGCQGFILDN YISALSAL WHPDCFVCRECFAPFSGGSFFEHE GRPLCENIFHARRGSLCATGLPVTGRCVSA LGRRRHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG GRPLCENIFHARRGSLCATGLCPVTGRCVSA LGRRRHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG FIVYNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEQLAGLAGLDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CLINATEWVREHAGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGQSACEEFWARLR KLNWKRILRIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTIYVHMYTRTCMHTYPYMYMNSV LISSELLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSIRKVNABPGCVVTNACAGSOCPPHAACRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	, , ,	2220	**	, 022	,	1402	
PAAPAAPPFSSSSGVLGTGLCELDRLLQELNA TQFNITDEIMSQFPSKVASGGKEDQSEDKK RPSLPSSPSPGLPKASATSATLEJDRLMASLSD FRVQNHLPASGPTQPPVVSSTNEGSPSPPPPTG KGSLDTMLGLLQSDLSRGVPTQAKGLCGSC NKPPJAGQVVTALGRA WHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHEHFCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAFFSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCRPLTKGSFQERAGKPY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIYYNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAFAQLAELDLLASMFFGENE LIVNDQLAVAELDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CLINATEWYREHASGYVSRDTSSPTTGSTVQ SVDLIFTRLWYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILRIREDFTDTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSELLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHNCEHRMD					•		1
TOFNITDEIMSOFPSSKVASGEQKEDQSEDKK RPSLPSSPSPGLPKASATSATLELDRLMASLSD FRVQNHLPASAFTSATLELDRLMASLSD FRVQNHLPASOFTOPPVVSSTNEGSPSPPETTG KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPIAGQVVTALGRA WHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCRDFLQLFAPRCOGCQGPILDN YISALSALWHPDCFVCRECFAPFSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCTLKLFG POPTICL LFG CQPCTLKLFG SQPCTLKLFG TIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG SLOPPSGLKQSSHLSSSWDFRHAPTHPET YTCKMEMEQAEAQLAELDLLASMFFGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWYREHASGYVSRDTSSSFTTGSTVQ SVDLIFTRLWYSHHIYNKCKRKNILEWAKEL SLSGFSMGKPGVVCVEGPQSACEFWARIR KLNWKRLIRIRRDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSELLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCOPGTYGFGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRND							
RPSLPSSPSGLPKASATSATLELDRLMASLSD FRVQNHLPASGPTOPPVVSSTNGGSPSPEPTG KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPIAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RIHKMYTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPPSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIYYNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPBT YTCYKNIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPBI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWYREHASGYVSKDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEFWARLR KLNWKRILRIIREDIPFDGTNDETERQKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVTTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQQVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD					,		
FRVONHLPASGPTOPPVVSSTNEGSPSPEPETG KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPIAGQVVTALGRAWHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RIKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCREPLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPFSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LJVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKNAMFSLACILPFKYPAVLPEI TVRSVLLSRQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLUNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD							
KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC NKPJAGQVVTALGRA WHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDDCFVCRECFAPFSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG YCKSQAWG YCKSQAWG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMGQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRIIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV LISSEILLIPSKYLFESK 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK GALTWSHPLLAVCPQQVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGFGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD					•		
NKPIAGQVVTALGRA WHPEHFVCGGCSTAL GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSAL WHPDCFVCRECFAPFSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPPHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKACIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEFFWARLR KLNWKRILIRIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGGCVTTFVVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGFGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD							KGSLDTMLGLLQSDLSRRGVPTQAKGLCGSC
RHKMVTALGTHWHPEHFCCVSCGEPFGDEG FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPFSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTIGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILRIIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD							NKPIAGQVVTALGRAWHPEHFVCGGCSTAL
FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN YISALSALWHPDCFVCRECFAPFSGGSFFEHE GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG 907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTIGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILRIIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHFLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGFGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD			1	1			GGSSFFEKDGAPFCPECYFERFSPRCGFCNQPI
907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG 908 2258 SLQPPSGLKQSSHLSLSSSWDFRHAPTHPET 910 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV 909 2259 A 7884 212 4874 GALTWSHPLLAVCPGVVLGSTPSGSPALLP 910 2260 A 7884 212 4874 GALTWSHPLCAVCAGPCVAGCPFHANCCDL 907 VTSALSALWHPDCFVCRECFAPFSGGSFFEHE 917 GRPLCENFTHARRSSLCATGLPVTGCVSA 908 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV 118SEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPGGVWLGSTPSGSPALLP 917 SALSALWHPDCFVCRCFAPFSGGSFFEHE 918 GRPLCENFTHARRSSLCATCGLPVTGRCVSA 919 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV 118SEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPGGVWLGSTPSGSPALLP 911 PSHRVNAEPGCVVTNACASGPCPPHANCRDL 912 VYGSTCQPGYYGPGCVDACLLNPCQNQG 913 SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD		1		٠ [
GRPLCENHFHARRGSLCATCGLPVTGRCVSA LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY CQPCFLKLFG QPCFLKLFG FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTIGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRIEDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLINPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD							FHEREGRPYCRRDFLQLFAPRCQGCQGPILDN
LGRRFHPDHFTCTFCLRPLTKGSFQERAGKPY COPCFLKLFG					ļ		
907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG 908 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG 908 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG 909 110 110 11172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG 909 110 110 11172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG 909 110 110 110 11172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG 909 110 110 110 11172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG 909 110 110 110 110 110 110 110 110 110					ŀ		GRPLCENHFHARRGSLCATCGLPVTGRCVSA
907 2257 A 7828 1792 1671 FIYVNQSFAPSPDQEVGTLYECFGSDGKLVLH YCKSQAWG 908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTIGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRIIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD							
908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSPTIGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRIIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	007	2257	- 	7000	1700	1/81	
908 2258 A 7842 110 1172 KLSCPCSHGTRVTAVRGPRLKAGVQWHDLG SLQPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSK VYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTIGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRIIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	<i>3</i> 01	2231	^	/026	1/92	10/1	
SLOPPPSGLKQSSHLSLSSSWDFRHAPTHPET YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTIGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRIREDIPFDGINDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	908	2258	_	7842	110	1172	
YTCPKMIEMEQAEAQLAELDLLASMFPGENE LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTIGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	700	٥٤عم	^	1042		11/2	
LIVNDQLAVAELKDCIEKKTMEGRSSKVYFTI NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTIGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRIREDIPFDGINDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	- 1	Ì	Į	}	,	ļ	
NMNLDVSDEKMAMFSLACILPFKYPAVLPEI TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTIGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRIIREDIPFDGTINDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD		1	İ	1			I IVNDOI AVARI VDOIEVVTARCDOOVIATED
TVRSVLLSRSQQTQLNTDLTAFLQKHCHGDV CILNATEWVREHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRIIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD					-	-	
CILNATEWVRÉHASGYVSRDTSSSPTTGSTVQ SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRIIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD			1		1		
SVDLIFTRLWIYSHHIYNKCKRKNILEWAKEL SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRIIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD			.	1	. !	ļ	
SLSGFSMPGKPGVVCVEGPQSACEEFWARLR KLNWKRILIRIIREDIPFDGTNDETERQRKFSIF EEKVFSVNGARGNHMDFGQLYQFLNTKGCG DVFQMFLWV 909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	1	ľ	ľ	ĺ	ŀ		SVDLIFTRI WIYSHHIYNKCKRKNII EWAKEI
KLNWKRILIRIIREDIPFDGTNDETERQRKFSIF		l	j	- 1			
909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	l	l	- 1				
909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	l	1	1		1	j	
909 2259 A 7870 3067 2923 EGICVYTFIYVHMYTRTCMHTYPYMYMNSV LISSEILLIPSKYLFESK 910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	ŀ		- 1	ļ		1	• • • • • • • • • • • • • • • • • • • •
910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	909	2259	Α	7870	3067	2923	
910 2260 A 7884 212 4874 GALTWSHPLLAVCPQGVWLGSTPSGSPALLP PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD				ļ			
PSHRVNAEPGCVVTNACASGPCPPHANCRDL WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	910	2260	A	7884	212	4874	
WQTFSCTCQPGYYGPGCVDACLLNPCQNQG SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	ı	ĺ	- 1	İ	ı	l	· ·
SCRHLPGAPHGYTCDCVGGYFGHHCEHRMD	l	l			I	į	
	j	-			1	į	

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						TNGQCHCKEFHYRPRGSDSCLPCDCYPVGST SRSCAPHSGQCPCRPGALGRQCNSCDSPFAEV TASGCRVLYDACPKSLRSGVWWPQTKFGVL ATVPCPRGALGLRGAGAAVRLCDEAQGWLE PDLFNCTSPAFRELSLLLDGLELNKTALDTME AKKLAQRLREVTGHTDHYFSQDVRVTARLL AHLLAFESHQQGFGLTATQDAHFNENLLWA GSALLAPETGDLWAALGQRAPGGSPGSAGLV RHLEEYAATLARNMELTYLNPMGLVTPNIML SIDRMEHPSSPRGARRYPRYHSNLFRGQDAW DPHTHVLLPSQSPRSPSEVLPTSSSIENSTTSS VVPPPAPPEPEPGISIIILLVYRTLGGLLPAQFQ AERGARLPQNPVMNSPVVSVAVPHGRNFLR GILESPISLEFRLLQTANRSKAICVQWDPPGLA EQHGVWTARDCELVHRNGSHARCRCSRTGT FGVLMDASPRERLEGDLELLAVFTHVVVAVS VAALVLTAAILLSLRSLKSNVRGIHANVAAA LGVAELLFLLGIHRTHNQLVCTAVVILLHYFF LSTFAWLFVQGLHLYRMQVEPRNVDRGAMR FYHALGWGVPAVLLGLAVGLDPEGYGNPDF CWISVHEPLIWSFAGPVVLVIVMNGTMFLLA ARTSCSTGQREAKKTSALTLRSSFLLLLLVSA SWLFGLLAVNHSILAFHYLHAGLCGLQGLAV LLLFCVLNADARAAWMPACLGRKAAPEEAR PAPGLGPGAYNNTALFEESGLIRITLGASTVSS VSSARSGRTQDQDSQRGRSYLRDNVLVRHGS AADHTDHSLQAHAGPTDLDVAMFHRDAGA DSDSDSDLSLEEERSLSIPSSESEDNGRTRGRF QRPLCRAAQSERLLTHPKDVDGNDLLSYWPA LGECEAAPCALQTWGSERRLGLDTSKDAAN NNQPDPALTSGDETSLGRAQRQRKGILKNRL QYPLVPQTRGAPELSWCRAATLGHRAVPAAS YGRIYAGGGTGSLSQPASRYSSREQLDLLLRR QLSRERLEEAPAPVLRPLSRPGSQECMDAAPG RLEPKDRGSTLPRRQPPRDYPGAMAGRFGSR DALDLGAPREWLSTLPPPRTTRDLDPQPPPLP LSPQRQLSRDPLLPSRPLDSLSRSSNSREQLDQ VPSRHPSREALGPLPQLLRAREDSVSGPSHGP STEQLDILSSILASFNSSALSSVQSSSTPLGPHT TATPSATASVLGPSTPRSATSHSISELSPDSEPR DTQALLSATQAMDLRRRDYHMERPLLNQEH LEELGRWGSAPRTHQWRTWLQCSRARAYAL LLQHLPVLVWLPRYPVRDWLLGDLLSGLSVA IMQLPQGLAYALLAGLPPVGTYSFYPVFIY FLFGTSRHISVESLCVPGPVDT
911	2261	A	7890	21	806	EFGTSRSSRSMAEDLGLSFGETASVEMLPEHG SCRPKARSSSARWALTCCLVLLPFLAGLTTYL LVSQLRAQGEACVQFQALKGQEFAPSHQQV YAPLRADGDKPRAHLTVVRQTIPTQHFKNQFP ALHWEHELGLAFTKNRMNYTNKFLLIPESGD YFIYSQVTFRGMTSECSEIRQAGRPNKPDSITV VITKVTDSYPEPTQLLMGTKSVCEVGSNWFQ PIYLGAMFSLQEGDKLMVNVSDISLVDYTKE DKTFFGAFLL
912	2262	A	7891	1263	111	ACGIRHEGALPGLTATPEAMLRFLPDLAFSFL LILALGQAVQFQEYVFLQFLGLDKAPSPQKFQ PVPYILKKIFQDREAAATTGVSRDLCYVKELG VRGNVLRFLPDQGFFLYPKKISQASSCLQKLL YFNLSAIKEREQLTLAQLGLDLGPNSYYNLGP ELELALFLVQEPHVWGQTTPKPGKMFVLRSV

CEO ID	SEQ ID	Met	SEQ	Predicted	Dendisted and	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID NO: of	NO: of	hod	ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
l				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
1	1		1	peptide	ł	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
			1			PWPQGAVHFNLLDVAKDWNDNPRKNFGLFL
1	Į.	j			1	EILVKEDRDSGVNFQPEDTCARLRCSLHASLL
						VVTLNPDQCHPSRKRRAAIPVPKLSCKNLCH
						RHQLFINFRDLGWHKWIIAPKGFMANYCHGE
		ĺ	1		1	CPFSLTISLNSSNYAFMQALMIHAVDPEIPQAV CIPTKLSPISMLYQDNNDNVILRHYEDMVVD
	ſ	1	1			ECGCG
913	2263	A	7892	15	849	ASRLPRGPGCGADMRPLLGLLLVFAGCTFAL
1		١			***	YLLSTRLPRGRRLGSTEEAGGRSLWFPSDLAE
1	1	ł	l		1	LRELSEVLREYRKEHQAYVFLLFCGAYLYKQ
						GFAIPGSSFLNVLAGALFGPWLGLLLCCVLTS
						VGATCCYLLSSIFGKQLVVSYFPDKVALLQR
ļ			1			KVEENRNSLFFFLLFLRLFPMTPNWFLNLSAPI
1	!				[LNIPIVQFFFSVLIGLIPYNFICVQTGSILSTLTS
						LDALFSWDTVFKLLAIAMVALIPGTLIKKFSQ
01.4	20/4		2000			KHLQLNETSTANHIHSRKDT
914	2264	Α	7893	815	959	KSGWVWWLTPLIPALWEAQTEGSLRPEVKN
915	2265	A	7909	3	641	RLSNITRPFFSKKKKILV
313	2203	^	עטפו	,	041	HASGPGGLLRRRRGSGANMPVARSWVCRKT YVTPRRPFEKSRLDQELKLIGEYGLRNKREV
						WRVKFTLAKIRKAARELLTLDEKDPRRLFEG
1	ļ					NALLRRLVRIGVLDEGKMKLDYILGLKIEDFL
1						ERRLQTQVFKLGLAKSIHHAHVLIQQCHIRVR
						EQVVNILFFTVRLDSQKHIDFSLCFPIGVANPS
						HVKRKNASKGQGGAGARDDEEEE
916	2266	Α	7914	3	967	VAHTQWHTCQRLSQLTHRSILKYLLIDTHAC
				'		QVLILKHTHASLSLPSCQECFPSSIPSASHMVS
						HPHPPPSPRWGQTPEGLPAASPCGPGPRSCFS
				•		SILPTGDSWGMLACLCTVLWHLPAVPALNRT
						GDPGPGPSIQKTYDLTRYLEHQLRSLAGTYLN
				1		YLGPPFNEPDFNPPRLGAETLPRATVDLEVW RSLNDKLRLTQNYEAYSHLLCYLRGLNRQAA
	1			l		TAELRRSLAHFCTSLQGLLGSIAGVMAALGY
						PLPQPLPGTEPTWTPGPAHSDFLQKMDDFWL
						LKELQTWLWRSAKDFNRLKKKMQPPAAAVT
						LHLGAHGF
917	2267	A	7921	2	1166	RPRRGQGLVQEVQTENVTVAEGGVAEITCRL
						HQYDGSIVVIQNPARQTLFFNGTRALKDERFQ
						LEEFSPRRVRIRLSDARLEDEGGYFCQLYTED
				}		THHQIATLTVLVAPENPVVEVREQAVEGGEV
						ELSCLVPRSRPAATLRWYRDRKELKGVSSSQ
						ENGKVWSVASTVRFRVDRKDDGGIICEAQN
						QALPSGHSKQTQYVLDVQYSPTARIHASQAV VREGDTLVLTCAVTGNPRPNQIRWNRGNESL
						PERAEAVGETLTLPGLVSADNGTYTCEASNK
						HGHARALYVLVVYGESRLRPTEGGGGAPDP
						GAVVEAQTSVPYAIVGGILALLVFLIICVLVG
						MVWCSVRQKGSYLTHEASGLDEQGEAREAF
						LNGSDGHKRKEEFFI
918	2268	A	7938	3	2653	RRRLPPASPPSSSVSSSLSPSAVVMACRWSTK
						ESPRWRSALLLLFLAGVYGNGALAEHSENVH
						ISGVSTACGETPEQIRAPSGIITSPGWPSEYPAK
						INCSWFIRANPGEIITISFQDFDIQGSRRCNLD
						WLTIETYKNIESYRACGSTIPPPYISSQDHIWIR
						FHSDDNISRKGFRLAYFSGKSEEPNCACDQFR
						CGNGKCIPEAWKCNNMDECGDRSDEEICAKE
						ANPPTAAAFQPCAYNQFQCLSRFTKVYTCLP
						ESLKCDGNIDCLDLGDEIDCDVPTCGQWLKY
	ليبسيا		<u></u> j			FYGTFNSPNYPDFYPPGSNCTWLIDTGDHRK

	T 44		Lanc			7
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence]]	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			i	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
l		ł		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
I		l		peptide	•	/=possible nucleotide deletion, \=possible
	1			sequence		nucleotide insertion
	 			boqueiioc		VILRFIDFKLDGTGYGDYVKIYDGLEENPHK
l			ļ	1		LLRVLTAFDSHAPLTVVSSSGQIRVHFCADKV
					ļ	NAARGFNATYQVDGFCLPWEIPCGGNWGCY
	1	1		l .		
	i					TEQQRCDGYWHCPNGRDETNCTMCQKEEFP
	į			1		CSRNGVCYPRSDRCNYQNHCPNGSDEKNCFF
			ł			CQPGNFHCKNNRCVFESWVCDSQDDCGDGS
		i	Ì		<u></u>	DEENCPVIVPTRVITAAVIGSLICGLLLVIALG
						CTCKLYSLRMFERRSFETQLSRVEAELLRREA
	1	1			ļ	PPSYGQLIAQGLIPPVEDFPVCSPNQASVLENL
	ļ			}		RLAVRSQLGFTSVRLPMAGRSSNIWNRIFNFA
	ļ	1				RSRHSGSLALVSADGDEVVPSQSTSREPERNH
1	1	1	1	[[THRSLFSVESDDTDTENERRDMAGASGGVAA
				1		PLPQKVPPTTAVEATVGACASSSTQSTRGGH
		ł		i		ADNGRDVTSVEPPSVSPARHQLTSALSRMTQ
		1		ļ		GLRWVRFTLGRSSSLSQNQSPLRQLDNGVSG
l	ł	ł				REDDDDVEMLIPISDGSSDFDVNDCSRPLLDL
						ASDQGQGLRQPYNATNPGVRPSNRDGPCERC
	i	}		İ		GIVHTAQIPDTCLEVTLKNETSDDEALLLC
919	2269	A	7951	1674	1839	VVRVTCCPPARSTTERTNAYDEEDCVEMVAS
717	2209	^	1931	1074	1037	
	0070		5052	45	670	GGWNDVACHTTMYFMCEFDKKNM
920	2270	Α	7953	47	572	GGRASWPEQAKEPRREGHTDKQQTEDVLAA
						GLRCLPHLPAICARRMSPAFRAMDVEPRAKG
	ł	1		ŀ		VLLEPFVHQVGGHSCVLRFNETTLCKPLVPRE
		1				HQFYETLPAEMRKFTPQYKGKSQLLEGLPHW
ļ	ļ	l		1	ĺ	RGDVRDRGHGRPWQPSLEPSLPPTLCFPSLSS
	1				<u></u>	FSSSWPSAQHLTPSVFNPW
921	2271	Α	7957	612	812	RSGRTVVTGIGYSKALQSSNRNTKSLLQNEF
		ļ				MMVYSFRALSFKESTWATFQHGGEATKSRSL
}	1			١ .		SSTQ
922	2272	A	7967	1443	1660	ENITEKWKEIWMCRGNKKSCCWTFIKDRHLT
						VSCCKSKSGETLLICIFCSNLVGFFFFGIRGFSN
						WELVKPN
923	2273	A	7981	1	3023	GSAPRAATAMARARPPPPPSPPPGLLPLLPPLL
/23	22,3	1 11	1701	1	3020	LLPLLLLPAGCRALEETLMDTKWVTSELAWT
	1	1	1			SHPESGWEEVSGYDEAMNPIRTYOVCNVRES
1	1	l		}		
	1			ĺ		SQNNWLRTGFIWRRDVQRVYVELKFTVRDC
Į		l		1		NSIPNIPGSCKETFNLFYYEADSDVASASSPFW
		1	1	1		MENPYVKVDTIAPDESFSRLDAGRVNTKVRS
1	1		i	ſ		FGPLSKAGFYLAFQDQGACMSLISVRAFYKK
1		1	1			CASTTAGFALFPETLTGAEPTSLVIAPGTCIPN
1			1	1		AVEVSVPLKLYCNGDGEWMVPVGACTCATG
						HEPAAKESQCRPCPPGSYKAKQGEGPCLPCPP
1			1	ļ		NSRTTSPAASICTCHNNFYRADSDSADSACTT
	1			<u>[</u>		VPSPPRGVISNVNETSLILEWSEPRDLGVRDD
ľ	ľ	Ι .	l			LLYNVICKKCHGAGGASACSRCDDNVEFVPR
				[QLGLSEPRVHTSHLLAHTRYTFEVQAVNGVS
1				l		GKSPLPPRYAAVNITTNQAAPSEVPTLRLHSS
		1	1	1		SGSSLTLSWAPPERPNGVILDYEMKYFEKSEG
1	1	1	ł	l		IASTVTSOMNSVOLDGLRPDARYVVOVRART
				1		VAGYGOYSRPAEFETTSERGSGAOOLOEOLP
	1]		LIVGSATAGLVFVVAVVVIAIVCLRKQRHGS
l			l			DSEYTEKLQQYIAPGMKVYIDPFTYEDPNEA
j	J	1		J]	VREFAKEIDVSCVKIEEVIGAGEFGEVCRGRL
	}	l	1	ļ		
1	†			İ		KQPGRREVFVAIKTLKVGYTERQRRDFLSEA
1		1	l			SIMGQFDHPNIIRLEGVVTKSRPVMILTEFME
						NCALDSFLRLNDGQFTVIQLVGMLRGIAAGM
	1					KYLSEMNYVHRDLAARNILVNSNLVCKVSDF
	1			1		GLSRFLEDDPSDPTYTSSLGGKIPIRWTAPEAI
L_	L	<u> </u>	L	<u> </u>	<u></u>	AYRKFTSASDVWSYGIVMWEVMSYGERPY
						

NO. of NO. of No. of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine.
Decidide Decidide Decidion Decidide Decidion Decidide Decidion Decidide							D=Aspartic Acid, E=Glutamic Acid.
Sociation Soci	nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
Sequence	cotide	seq-	Ì	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
mino acid reduce peptide sequence ide pertia peptide peptide pertia pert		цепсе					M=Methionine, N=Asparagine, P=Proline,
	uence]	914			Q=Glutamine, R=Arginine, S=Serine,
Poptide Sequence	[1	Ì			
		1		1		sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
WDMSNQDVINAYEQDYRLPPMOCPTALEQ				ł		ļ	
LMLDCWYRDNALRYKS9QNYNTLDKLIRNAA SLKVIASAQSGMSQPLLDRYPDYTTFTYOD WLDAIKMGRYKESFVSAGFASPDLVAQMTA EDLLRIQVITAGHQKKLISSIQDMRLQMNOT LPVQV	ļ		ļ	ļ	sequence	1	
SILKVIASAQSGMSQPILIDRIYPDYTIFTITUDG WLDAIKMGRYKESFYSAGASSPIUAQMTA EDILBIGVILAGHQKKLISSIQDMRLQMNQT LPVQV	l	1	l	1		ĺ	WDMSNQDVINAVEQDYRLPPPMDCPTALHQ
WLDAIKMGRYKESFVSAGFASFDLVAQMTA EDLLRIGVTLAGHCKLISSIQDMRLQMNOT LPVQV						ĺ	LMLDCWVRDRNLRPKFSQIVNTLDKLIRNAA
				'			SLKVIASAQSGMSQPLLDRTVPDYTTFTTVGD
1. 1. 1. 1. 1. 1. 1. 1.		1					
924 2274 A 7985 1 503 FRPRTIKATAMYLEHYLDSIENLECELQRING	l	}	1	1			
QLAMRELDORTEDKKAEIDILAAEYISTVKTLS	924	2274	_	7085		502	
PDQRVERLQKIQNAYSKCKEYSDDKVQLAKAMEG] /27	22/4	^	7703	1	303	OLAREL DODTEDVY A FIDU A A FYICT WAY O
925 2275 A 7994 447 589]		ļ	DUODAEDI UAIONY AGRAMANINI AM
SDFESSGGRGLKKGRGQKEKRGSRGRGRTS			}				
P25 2275							SDEESSGGRGI KKGRGOKEKRGSPGPGPGPTS
925 2275 A 7994 447 589			ł	1			
SPIRCYCOHWPHCVHC	925	2275	A	7994	447	589	
926	,]			
	926	2276	A	7996	925	582	
927 2277 A 7998 2 353 RIQRPLNSRSPNIBLFVKABLTAKQATMKLSV CLLLVTLALCCYQANAECPCHAUSELLDFFFI SPEJFKLSLAKFDAPPEAVAAKLGVKRCTDQ MSLQKRSLIAEVLVKILKKCSV CLLLVTLALCCYQANAECPCHAUSELLDFFFI SPEJFKLSLAKFDAPPEAVAAKLGVKRCTDQ MSLQKRSLIAEVLVKILKKCSV MSLQKRSLIAEVLVKILKKCSV LAPPROQPGTRTQPRSHPAANDFSAMSAAQ ARGIRATYHRLLDKVEIM.PEKLRPLYNHPA GPRTVFFWAPIMKWGLVCAGLADMARPAEK LSTAQSAVLMATOFIWRNWSYSLVIIPKNWSLFA VNFFVQAAGASQLFRIWRYNQELKAKAHK KSTAQSAVLMATOFIYTAGPRI.SASASKELLM KLRRATGSSPOPHITTYAGPRI.SASASKELLM KLRRATGYSFVNCKKALETCGGOLKQAEIWL HKAQKEGWSKAAKLQGRKTKEGLIGLLQE GNTTVLVEVNCETIDFVSRNLKFQLLVQQVAL GTMMHCQTILKDQPSAYSKGFLNSEI.GI.PA GPDRGSI.KDQLAI.GKGEMMILKRAAW KVPSGFYVGSYVHGAMQSPSLHKLVLQKVG ALVICETSBCKTRLEDVGRRLIGQHVCMAPG GPDRGSI.KDQLAI.GKGEMMILKRAAW KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSBCKTRLEDVGRRLIGQHVCMAPG GPDRGSI.KDQLAI.GKGEMMILKRAAW KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSBCKTRLEDVGRRLIGQHVCMAPG GPDRGSI.KDQLAI.GKGEMMILKRAAW KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSBCKTTLEDVGRRLIGQHVCMAPG GPDRGSI.KDQLAI.GKGEMMILKRAAW KVPSGFYVGSYVHGAMGSPSLHKLVLGKYG ALVICETSBCKTTLEDVGRRLIGGHVCMAPG GPDRGSI.KDQLAI.GLGENGHMILKRAAW KVPSGFYVGSYVHGARGSSFI.YGYN LSVVNADPPYKAFYBERGRGSFF.YGYN LSVVNADPPYKAFYBERGRGSFF.YGYN LSVVNADPPYKAFYBERGRGSFF.YGYN LSVVNADPPYKAFYBERGRGSFF.YGYN LSVVNADPPYKAFYBERGRGSFF.YGYN LSVVNADPPYKAFYBERGRGSFF.YGYN LSVVNADPPYKAFYBERGRGSFF.YGYN LSVVNADPPYKAFYBERGRGSFF.YGYN GRAWFYTNSIFGKAGMPAKFYBERGRGSG GPTLTTLTLAQHAPWYRQVVVIVITMACYQL CGIANIWYTNSIFGKAGMPAKFYPVLXSTG GETLAAVFSGLVIEHLGREFTVINTLSTG GETLAAVFSGLVIEHLGRPPLLIGGGGLMGLFF GTLTTLATLAQHAPWAVQVVILSTGFGGGLMGLFF GTLTTLATLAQHAPWAVQVVILSTGGGGLMGLFF GTLTTLATLAQHAPWAVQVVILSTGGGGLMGLFF GTLTTLATLAQHAPWAVQVVILSTGGGGLMGLFF GTLTTLATLAQHAPWAVQVVILSTGGGGLMGLFF GTLTTLATLAQHAPWAVQVVILSTGGGGLMGLFF GTLTTLATLAQHAPWAVGAVAPFACCSHAFTAGTTOAYYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEK DSAVTDGKINGRP AAGAVVSAMPKAKGKTRQKFGYSVNKKKL NRNARRKAAPRECSHRIAWDHAKSVQNIL							
927 2277 A 7998 2 353 RIQRPLNSRSPNHSLFVKAELTAKQATMKLSV CLLVTLALCCYQANAEPCPALVSELLDFFFI SEPLFKLSLAKFDAPPEAVAKLGVKRCTDQ MSLQKRSSLAEPLVKLILKKCSV 928 2278 A 8004 130 588 LAPLRCQPGTRTQPRSHPAANDFSAAMSAAG ARGLRATYHRLDKVELMLPEKLRPLYNHSH GPRTVYFWAPIMKWGLVCAGLADMARPAEK LSTAQSAVLMATOFHWSRYSLVIPKNWSLFA VNFFVQAAGASQLFRIWRYNQELKAKAHK LSTAQSAVLMATOFHWSRYSLVIPKNWSLFA VNFFVQAAGASQLFRIWRYNQELKAKAHK KLRKTGYSFYACKAGLEGHLKQAEHWL HKEAQKEGWSKAAKLQGRKTKEGLIGLQE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMHHCOTILKDPSAVSGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKVY ALVICETSEGKTNLEDDVGRLGQHVVGMAPL SVGSLDDEPGGBETKMLSQPYLDPSITLGQ YQVQQOVSVVDFVRFECGEGEAAETE GLVPLTDDTSHAGPPGGRALHEDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTFYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIKMIGKVLGRK HTTLLANNGFAISAALLMACSLQAGAFEMLIV GRFMGIDGGVALSVLPMYLSEISKERGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWYPLF GVIVVPAVVQLLSLFPLDSFRYLLLEKHNEA RAVKAFQIFLGKADVSQEVEEVLAESRVQRS IRVSVLEULRAPVYRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVUEHLGRPLLIGGGGLMGLFF GTLTITLTLQDHAPWYSIVSIGGILAISFCSG PGGPFILTGEFPQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATAICTIGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEK DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRQKFGYSVNRKRL							
927						•	RDTQPILPLGGRYYITIRO
CLLLVTLALCCYQANAEFCPALVSELLDFFF SEPLFKLSLAKFDAPPEAVAAKLGVKRCTDQ SEPLFKLSLAKFDAPPEAVAAKLGVKRCTDQ MSLQKRSLIAEVLYKLIKKCSV MSLQKRSLIAEVLYKLIKKCSV MSLQKRSLIAEVLYKLIKKCSV A	927	2277	Α	7998	2	353	RIORPLNSRSPNHSLFVKAELTAKOATMKLSV
MSLQKRSLIAEVLVKILKKCSV							
MSLQKRSLIAEVLVKILKKCSV							
929 2279 A 8007 2 1016 EFARRYIAREMSLIKATION OF AGAING AGASOL FRIWTYNOELKAKAHK PSEP 2279 A 8007 2 1016 EFARRYIAREMSLIKSIKVILVARTOSYP AGSLIRQSPOPRHITYYAGPRISASASIKELLM KLARKTGYSFYNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQGRITKEGLIGLLOE GNITVLVEVNCETDFYSRNLKFQLLVQQVAL GTMMHQTILKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDUGRRIGOHVVGMAPL SVGSLDDEPGGEAETKMLSOPYLLDPSITLGQ YVQPGVSVVDFVRFCGEGEEAAEITE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYTKAFYNESWERHGRPDPDTL TILWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKAUVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG ETILAVFSGLVIEHLGRPPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFLITGEFFQQSGRAAFIIAGTVNWLSN FAVGLLFFFIQSKLDTYCFLVFAGTICTIGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL							
929 2279 A 8007 2 1016 EFARRVFIAAREMSLLRSLRVFLVARTGSYP AGSLLROSPOPRHITYAOPILSAKASKELLM KLRRKTGYSFYOKKALETCGGDLKQAEIW HKEAQKEGWSKAAKLQGRKTKEGLIGLOE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKQDPSAYSKGFLNSSELSGLPA GPPREGSLKDQLALAIGKLGENMILKRAAW KVPSGFYVGSVYHGAMQSPSLHKLUJGKYG ALVICETSEQKTNLEDVGRRLOGHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLDPSITLGQ YVQPQGVSVVDFVRFECGEGEEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLYVTDDTSHAGPPGGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERHGRPDPDTL TILWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAFIFICIGFYTGQLLCIPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELRAPYVRWQVTVTVTMACYQL CGLNAIWFYTNSIFGKAGPPAKIPYVTLSTGG IETLAAVFSGLVTEHLIGGRPAKIPYVTLSTGG IETLAAVFSGLVTEHLIGRSPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGATYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	928	2278	Α	8004	130	588	
STAQSAVLMATGFIWSRYSLVIIPKNWSLFA VNFFVQAGASQLFRIWSRYSLVIIPKNWSLFA VNFFVQAGASQLFRIWRYNQELKAKAHK 929 2279 A 8007 2 1016 EFARRVFIAAREMSLLRSLRVFLVARTGSYP AGSLLRQSPQPRHTFYAGPRLSASASSKELLM KLRKTGYSFVNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGLQE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTILKDQFSAYSKGFLNSSELSGLPA GPDREGSLKQALALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRILGQPHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFCCGEGEEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPTLDTTSHAGPGBGRAEATE GLVPTLDTTSHAGPGBGRAEATE GLVPTLDTTSHAGPGBGRAEATE GLVPTLDTTSHAGPGBGRAEATE GLVPTLDTTSHAGPGBGRAET GTLTLLANNGFAISAALLMACSLQAGAFGSSELYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TILLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTILLANNGFAISAALLMACSLQAGAFGMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFIGGLLGLEELLGKESTWPYLF GVIVVPAVVQLLSLFPLPDSPRYLLLEKHNEA RAVKAFQTIFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGREPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFLTGGFTQQSQRAAFIIAGTVNWLSN FAVGLLPFPIQKSLDTYCFLVPATICTIGATYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRI NRNARRKAAPRIECSHIRHAWDHAKSVRQNL							
929 2279 A 8007 2 1016 EFARRRVFIAAREMSLRSLRVFLVARTGSYP RAGSLRQSPQRHTFYAGPRLSASASKELLM KLRKTGYSFVNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGGHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEAAETE GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFIGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASSCSG PGGIPFLTGEFFQQSQRPAAFILAGTVNWLSN FAVGLLFFFIQKSLDTYCFLVFATICTIGATYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL							GPRTVFFWAPIMKWGLVCAGLADMARPAEK
930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRAGNENLECGHLING GLVPLTDDTSHAGPPGRAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLAGASSKELLE GLVPLTDDTSHAGPPGRAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRAGAGSSFLYGYN LSVVNAPTPYKAFYNESWERRHGRPDPDTL 11 LUWSVTVSIFAIGGLVGLALAGAGGGGEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALECDHLRSGV PGGRRKDWSCSLLVASLAGAGGSFLYGYN LSVVNAPTPYKAFYNESWERRHGRPDPDTL 12 LUWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HITLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELRAPPVRWQVTTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVPSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSVILAIIASFCSG PGGFPFLTGFFFQQSQRPAAFILAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL				1			
AGSLLRQSPQPRHTTYAGPRLSASASSKELLM KLRRKTGYSFVNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE GNTTVLVEVNCETDFYSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLAAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLQQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVFFCGGEBAAETE 930 2280 A 8008 3 1679 NSRVWGPVTFPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TILWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISFKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSFRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGGNAIWFYTNSIFGKAGPPAKIPYVTLSTFGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICTIGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRV 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	000						
SURRKTGYSFVNCKKALETCGGDLKQAEIWL HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGPIDPDTL TILLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTILLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTIGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVTVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRPFLLIGGFGLMGLFF GTLITITLTQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	929	2279	A	8007	2.	1016	EFARRRVFIAAREMSLLRSLRVFLVARTGSYP
HKEAQKEGWSKAAKLQGRITKEGLIGLLQE GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRLQQHVVQMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEAAETE GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISALLMASCLQAGAFEMLIV GRFIMGDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTIGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVTVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRPLLIGGFGLMGLFF GTLITITLTQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL					-		AGSLLRQSPQPRHTFYAGPRLSASASSKELLM
GNTTVLVEVNCETDFVSRNLKFQLLVQQVAL GTMMHCQTILKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTIGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGREPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL							KLRRKTGYSFVNCKKALETCGGDLKQAEIWL
GTMMHCQTLKDQPSAYSKGFLNSSELSGLPA GPDREGSLKDQLALAIGKI.GENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITI.GQ YVQPQGVSVVDFVRFECGEGEEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TILLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAHFICIGVFIGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSFRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	j	j					HKEAQKEGWSKAAKLQGRKTKEGLIGLLQE
GPDREGSLKDQLALAIGKLGENMILKRAAWV KVPSGFYVGSYVHGAMQSPSLHKLVLGKYG ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSOPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFIGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSFRYLLIEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			'				GNITVLVEVNCETDFVSRNLKFQLLVQQVAL
SVPSGFYVGSYVHGAMQSPSLHKLVLGKYG		-	-				
ALVICETSEQKTNLEDVGRRLGQHVVGMAPL SVGSLDDEPGGEAETKMLSQPYLLDPSITLGQ YVQPQGVSVVDFVRFECGEGEEAAETE 930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTIGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICTIGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		-					
930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFIGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	İ						
930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLFFPIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	1	ľ		' I			
930 2280 A 8008 3 1679 NSRVWGPWTEPSAGSLRPMARKQNRNSKEL GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			,	ł			
GLVPLTDDTSHAGPPGPGRALLECDHLRSGV PGGRRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	930	2280	A	8008	3	1679	
PGGRRKDWSCSLLVASLAGAFGSSFLYGYN LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TILLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRPPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL					-		
LSVVNAPTPYIKAFYNESWERRHGRPIDPDTL TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFIGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		ı	1		1		PGGRRKDWSCSLLVASLAGAFGSSFI VOVN
TLLWSVTVSIFAIGGLVGTLIVKMIGKVLGRK HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL						•]	
HTLLANNGFAISAALLMACSLQAGAFEMLIV GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFITGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	ŀ	1	1		1	Ì	
GRFIMGIDGGVALSVLPMYLSEISPKEIRGSLG QVTAIFICIGVFTGQLLGLPELLGKESTWPYLF GVIVVPAVVQLISLPFLPDSFRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		i			1	ĺ	
QVTAIFICIGVFIGQLLGLPELLGKESTWPYLF GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	ł	ľ	- 1	l l	}	I	
GVIVVPAVVQLLSLPFLPDSPRYLLLEKHNEA RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		İ	l	ļ	l	1	
RAVKAFQTFLGKADVSQEVEEVLAESRVQRS IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPVLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	ľ	l	l	i	l		
IRLVSVLELLRAPYVRWQVVTVIVTMACYQL CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL		l		ŀ	j		
CGLNAIWFYTNSIFGKAGIPPAKIPYVTLSTGG IETLAAVFSGLVIEHLGRRPLLIGGFGLMGLFF GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	ì	1	j		ļ		IRLVSVLELLRAPYVRWOVVTVIVTMACYOL.
931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	İ		l	1	{		
GTLTITLTLQDHAPWVPYLSIVGILAIIASFCSG PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL				ļ	ŀ		
PGGIPFILTGEFFQQSQRPAAFIIAGTVNWLSN FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL				ļ	l	1	
FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI DSAVTDGKINGRP 931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL				1	ļ		
931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL	İ	. /	1	1	- 1	İ	FAVGLLFPFIQKSLDTYCFLVFATICITGAIYL
931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL			- 1	İ	i		YFVLPETKNRTYAEISQAFSKRNKAYPPEEKI
931 2281 A 8009 861 300 AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL NRNARRKAAPRIECSHIRHAWDHAKSVRQNL							DSAVTDGKINGRP
	931	2281	A	8009	861	300	AAGAVVSAMPKAKGKTRRQKFGYSVNRKRL
AEMGLAVDPNRAVPLRKRKVKAMEVDIEFR	- 1		- 1		ļ	1	
							AEMGLAVDPNRAVPLRKRKVKAMEVDIEER

NO. of No. of N	SEO ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
			1		1		Amino acid sequence (A=Ananine C=Cysteine,
Sequence 09/49	1	II .	1.00	4		E .	
Sequence			1	1			
1914 ng to first mino aeid residue of pepitide pepitide	1	1 -					
maino acid residue of peptide residue of peptide sequence peptide		301100					
Presiduc of peptide sequence	10000		}	714			
Popular Sequence							
	J	1	l			sequence	
PRELVRRYYLNDLEARASIPEKGNTI.SRI	j	i					
				 	sequence	 	
TRKQRSKINTYKRFYPAEWQDFLDSLQKRK MEVE						1	
MRVE	ł	1	ł		ŀ	1	
932 2282 A 8011 412 1	1	i				!	
DIRGSHTSSPQRQSQRYPGKETABYLRAGK QGRQIPICPWPPPPPPPSGPGGCRGFH SLEAKARHPASYREMRGKVKMRALRRAPA STRASSRQIPPCWPPPPPPPPSGPGGCRGFHG NELGACARHPASYREMRGKVKMRALRRAPA STRASSRQIPPK STRASS	022	2202		0011	410	ļ	
933 2283 A 8012 147 1077 PPVPPPSREMGKVKMRRALRRAPA STRASSROPNPK PPVPPASRSDMAQNLKDLAGRLPAGFRGMGT ALKALLLGAGAVAYOVRESVITVEGGHRAIFF NRIGGVQQDTILAEGLHFRIWFQYPITVING PSMYQRLGLDVERRVLFRKADFSLILDD VAITELSFSREYTAAVEAKQVAQQEAQRAQ PSMYQRLGLDVERRVLFRKADFSLILDD VAITELSFSREYTAAVEAKQVAQQEAQRAQ LVEKAKQEQRQKVQAGEGARAAMLGEAL SKNPOYKLRKIRRAAMSKTIATSQNRIVLTA DNLVLINLQDESFTRGSDSLIKGKK ANSQLIVRACHERALFAKADFSLILDD VAITELSFSREYTAAVEAKQVAQQEAQRAQ LVEKAKQEQRQKVQAGEGARAAMLGEAL SKNPOYKLRKIRAAQNISKTIATSQNRIVLTA DNLVLINLQDESFTRGSDSLIKGKK RICKFRELHLMRNEARKLINHQEVVEEDKRK KLPANWEAKKARLBWELKEEEKKECARAR EDYEKVKLLEISAEDAERWERKKKRKNPDLG FSDYAAAQLRQYHRILTKQKRPMUDLE KQIEKRKYSRRHYNDDADIDVINERNAK KHGEFFPTINSLLHGTHVPSTEEDRMVDLE KQIEKRKYSRRHYNDDADIDVINERNAK NKKAREFYGKYTAEIKQNLERGTAV LVSSTVNILTEKAPWNSLAWTVTSVYFLKFL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSLSRL QGGGTGSTGMMDSALTLLIGIGSHRHSL QGGGTGSTGMMDSALTLLIGIGSHRHSL QGGGTGSTGMMDSALTLLIGIGSHRHSL QGGGTGSTGMMDSALTLLIGIGSHRHSL QGGGTGSTGMMDSALTUTAYTRAAMGML VYDIT NRKSFDNIRNWRNIEBHASADVEKMILGKK QGGGTGSTGMMDSKTATTAYTRAAMGML VYDIT NRKSFDNIRNWRNIEBHASADVEKMILGKG QGGGTGSTGMTGMT QGGGTGSTGMTGMT QGGGTGSTGMTGMT QGGGTGSTGMT QGGTGTGMT QGGGTGSTGMT QGGG	932	2202	^	.0011	412	1	
SLEAKARIPASYREMIGKVKMRRALRRAPA STRASSRQIPME STRASSRQIPM	1		1	İ i		1	DIRGSHFSSPQRQRSQRVPGKETARVLRAGK
		ł					
933 2283 A 8012	1				İ		
ALKLLLGAGAVAYGVRESVFTVEGGHRAIFF NRIGGVQQDTILAEGLHFRIPWGYPITDIRA RPRKISSPTGSKDQMVNISLRVLSRPYAQEL PSMYQRLGLDYEERVLPSIVNEVLKSVVAKF NASQLTQTRAQVSLLIRRELFEKDFSLLIDD VAITELSFSREYTAAVEAKQVAQQEAQAGF LVEKAK QEGGKIVQAEGEAAKMLGEAL SKNPGYIKLRKIRAAQNISKTIATSQNRIVLTA DNI-VINLQDESFTRGSDSLIKGKK DVEKAK QEGGKIVQAEGEAAKMLGEAL SKNPGYIKLRKIRAAQNISKTIATSQNRIVLTA DNI-VINLQDESFTRGSDSLIKGKK SKPGYKLRKERALQNISKTIATSQNRIVLTA DNI-VINLQDESFTRGSDSLIKGKK REPANWEAKKAR.EWELLSEEREKKECAARG EDYEKVKLLEISAEDAERWERKKKRKNPDLG FSDYAAAQLRQYHRLTKQIKPDMETYERLRE KHGEFFFTINSILLHGTHVYSTEEIDRMVIDLE KQIEKRDKYSRRRPYNDDADIDYINERNAF NKKAERFYGKYTAEIKQNLERGTAV GGGTGSTGMRDSALTLLIGERGTAV 935 2285 A 8027 59 310 LVSSTYNLLTEKAPWNSLAWTVTSYYFLKFL QGGTGSTGMRDSALTLLIGERGTAV GGGTGSTGMRDSALTLLIGERGTAV 936 2286 A 8032 1 639 SGRECNMAKTYDYLFKLLIGDSGVGKTCVL FRRSEDAFNSTISTIGIDKRRITELDGKRIKLQ IWDTAGQERFRITITAYYRGAMGIMLYVDTI NEKSFDNIRNWINEHEHASADVEKMILGRKC DVDIKKQVSKERGEKLALDYGIKFMETSAK ANNIVENAFFTLARDIKAKMDWINTHEHASADVEKMILGRKC DVDIKKQVSKERGEKLALDYGIKFMETSAK ANNIVENAFFTLARDIKAKMDALSWULTVISLL PLLEAQIPLCANLVPYPTINATLDRITGKWYI ASAFRNEEYNKSVQEQAATFFYFTPVKTEDTIF LREYGTRQDQCTYNTTYLNVQRENGTISRVV GQGFFF HLLILRDTKTYMLAFDVDEKMW GLSVYADKPETKEQIGEFYFALDCLRIPKSD VVYTIDWKKDKCEPLEKQHEKERQEGES 939 2289 A 8055 12 1039 SSVAEFPERVQUSQPONWNTADDRITGKSQ APGPPPFIGILSINPLMGDAVSDWSPMFEAA HGHOLSIRNLISQOWAVNITADHYSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTTL FNACVSGSWDCVNLLLQHGASVQPESDLASS HEAARRGHVECVNSLIAYGGNIDHKISHLT PLYLACENQUSQPONWNTADHYSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTTL FNACVSGSWDCVNLLLQHGASVQPESDLASS HEAARRGHVECVNSLIAYGGNIDHKISHLT PLYLACENQUSACKAKLELEGADVNGKQG DSPLHAVARTASEELACLILMDFGADTOKAN AEGKRYEL VPPESPALQLE REFERPEPSIMOL CRIRKCFGIQQHIKITKLVLPEDLKQFILH L							
NRIGGYQQDTILAEGI.HFRIPWFQYPITYDIRA RPPKISSPTGSKDLQMYNISLRVLSRPNAQEL PSMYQRLGLDYERVLPSIVPVLKSVVAKT NASQLTQRAQVSLLIRRELTERAKDFSLILDD VATTELSFSEYTAAVEAKQVQEQQRAQRAQ LVEKAKQEQRQKIVQAEGGEAEAAKMLGEAL SKNPGYVELRSKRAQNISKTATSONRYYLTA DNLVINLQDESFTRGSDSLIKGKK SKNPGYVELRSKRAQNISKTATSONRYYLTA DNLVINLQDESFTRGSDSLIKGKK SKNPGYVELRSKRAQNISKTATSONRYYLTA DNLVINLQDESFTRGSDSLIKGKK SQFSLSQVLVDSAEEGSLAAAAELAAQKREQ RRKFRELHIMRIBARKLINGVVEEDKRI, KLPANWEAKKARLEWELKSEEKKKECAARG EDVEKVILLEISLEADDER WERKKKRKINPDLG EDVEKVILLEISLEADDER WERKKKRKINPDLG FSDYAAAQLRQVHRLTKQIKPDMETYERLE KLPANWEAKKARLEWELKSEEKKKECAARG FSDYAAAQLRQVHRLTKQIKPDMETYERLE KLGEROKYSRRRYNDDADIDYINENNAKF NKKAERFYGKYTAERKQNLERGTAV NKKAERFYGKYTAERKQNLERGTAV OGGGTGSTGMRDSALTLLGIGFSHRHSLSIRL SQHSSPAPMYSQTFHILVLG 936 2286 A 8032 1 639 SGRECNMAKTYDYLFKLLIGDSGVGKTCVL FFRSEDARNSTFISTIGIDFKIRTIELDGKRIKLQ IVDTAGQERFRITTTAYYRGAMGIMLVYDIT NEKSFINIRNWIRNIEEHASADVEKMILGNKC DVNDKRQVSKEGEKLALDYGIKFMETSAK ANINVENAFFTLARDIKAKMKLLEGNSPQG SNQGVKITPDQQKRSSFRCVLL 937 2287 A 8039 393 311 EETIHSENSYLEKYPISANATUTIA 938 2288 A 8052 675 1334 LIPPATSTAWLHVPPGLSMALSWVLTVLSLL PLLEAQIPLCANLVPVPITNATLDRITIGKWFT LAEYQTRQDQCTYNITYLNVQRENGTISRYV GQGHFAHLLIRDTKTYMALAFVNDEKNIT ASAFNREYNKSVQEQATFFYFTPKTEDTIF LREYQTRQDQCTYNITYLNVQRENGTISRYV GQGHFAHLLIRDTKTYMALAFVNDEKNIT GLSVYADKFETTKFOLGEFYEALDCLRIKKSD VYYTDWKKUKCPIEKKGPEKERKQEGES SVAEFPERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGMOSSKP APGPSPORICLLRIKNSD APGPSPORICLLRIKSD APGPSPORICLLRIKSD APGPSPORICLLRIKSD APGPSPORICLLRIKSD APGPSPORICLLRIKSD APGREWARTASEELACLLMDFGADTOQKN AEGKRYELVPSSPAQACVKKLLESGADVNQGKOQ DSPLHAVARTASEELACLLMDFGADTOQKN AEGKRYELVPSSPAQACVKKLLESGADVNQGKOQ DSPLHAVARTASEELACLLMDFGADTOQKN AEGKRYELVPSSPAQACVKKLLESGADVNQGKOQ DSPLHAVARTASEELACLLMDFGADTOQKN AEGKRYELVPSSPAQACVKKLLESGADVNQGKOQ DSPLHAVARTASEELACLLMDFGADTOQKN AEGKRYELVPSSPAQACVKKLESGADVNQGKOQ DSPLHAVARTASEELACLLMDFGADTOQKN AEGKRYEL	933	2283	Α	8012	147	1077	
PRKISSPTGSKDLQMYNISLRVLSVPAKQEL PSMYQRLGILDYERVLPSIVLSVVAKY NASQLITQRAQVSLLIRRELTERAKDFSLILDD VAITELSFSREYTAAVEAKQVAQQEAQROF LVEKAKQEQGOKIVQAEGEAAKMI.GEAL SKNPGVIKLRKIRAAQNISKITATSQNRYLTA DNLVIALQDESFTRGSDSLIKGKK DNLVIALQDESFTRGSDSLIKGKK DNLVIALQDESFTRGSDSLIKGKK SKPPGVIKLRKIRAAQNISKITATSQNRYLTA DNLVIALQDESFTRGSDSLIKGKK PSMYANQLRQHILTKQIKEEKKKECARG RLRKFRELHLMRNEARKLNHQEVVEEDKRI KLPANWEAKKARLEWELKEEKKKECARG EDYEKVKLLEISAEDAERWERKKKKRNPDLG FSDYAAAQLRQYHRLTKQIKDHYTERINAK KLPANWEAKKARLEWELKEEKKKECARG EDYEKVKLLEISAEDAERWERKKKKRNPDLG FSDYAAAQLRQYHRLTKQIKDHYTERINAK KLPANWEAKKARLEWELKEEKKKECARG EDYEKVKLLEISAEDAERWERKKKKRNPDLG FSDYAAAQLRQYHRLTKQIKDHYTERINAK KLPANWEAKKARLEWELKEEKKKECARG EDYEKVKLLEISAEDAERWERKKKKRNPDLG FSDYAAAQLRQYHRLTKQIKDHYTERINAK KLPANWEAKKARLEWELKEEKKECARG EDYEKVKLLEISAEDAERWERKKKKRNPDLG FSDYAAAQLRQYHRLTKQIKDHYTKEN KHGEFFFTISNSLLHGTHYTERIDAKRWIPLK GUGGTOSTGMRSATLLIGGFSHRHSLSRL GUGGTOSTGMRSATLLIGGFSHRHSLSRL GUGGTOSTGMRSATLLIGGFSHRHSLSRL GUGGTOSTGMRSATLLIGGFSHRHSLSRL GUGGTOSTGMRSATLLIGGFSHRHSLSRL GUGGTOSTGMRSATLLIGGFSHRHSLSRL GUGGTOSTGMRSATLLIGGFSHRHSLSRL GUGGTOSTGMRSATLLIGGFSHRHSLSRL GUGGTOSTGMRSATLLIGGFSHRHSLSRL GUGGTOSTGMRSATLITITA THE STATEMEN WITH THE STATEMEN GUGGTOSTGMRST GUGG							
PSMYQRLGLDYEERVLPSIVNEVLKSVVAKF NASQLTIQRAQVSLLIRER IERAKDFSLILDD	1	1	1			1	
NASQLITQRAQVISLIERELTERAKDFSLILDD	1		1			!	RPRKISSPTGSKDLQMVNISLRVLSRPNAQEL
VAITELSPSREYTAAVEAKQVAQQEAGRAQF LVEKAKQEQRQKIVQAEGEAAKMLGEAL SKNPGYIKLRKIRAAQNISKITATSQNRYTTA DNLVLNLQDESFTRGSDSLIKGKK SKPGYIKLRKIRAAQNISKITATSQNRYTTA DNLVLNLQDESFTRGSDSLIKGKK SKPGSLSQVLVDSAEGESLAAAAELIAAQKREQ RLRKFRELHLMRNEARKLNHQEVVEEDKRL KLPANWEAKKARLEWELKEEGEKKECAARG EDYEKYKLLEISAEDAERWERKKRKNPDLG FSDYAAAQLRQYHRLTKQIKPDMETYERLRR KHGEEFFTSNSLLHGTHVPSTEEIDRNVIDLE KQIEKRDKYSRRPYNDDADDIDYINRRNAKF KKGEEFFTSNSLLHGTHVPSTEEIDRNVIDLE KQIEKRDKYSRRPYNDDADDIDYINRRNAKF KKAERFYGKYTAEIKQNLERGTAV 935 2285 A 8027 59 310 LVSSTVNLLTEKAPWNSLAWTVTSYVFLKFL QGGGTGSTGMRDSALTLLGIGFSIRRHSLSRL SQHSSPAPMYSQTFHILLV SQHSSPAPMYSQTFHILLV GGGTGSTGMRDSALTLLGIGFSIRRHSLSRL SQHSSPAPMYSQTFHILLV SQHSSPAPMYSQTFHILV SQHSSPAPMYSQTFHILV SQHSSPAPMYSQTFHILLV SQHSSPAPMYSQTFHILV SQHSSPAPMYSQTFHILV SQHSSPAPMYSQTFHILLV SQHSSPAPMYSQTFHILV SQHSSPAPMYSQTFHILV SQHSSPAPMYSQTFHILLV SQHSSPAPMYSQTFHILV SQH]			
1			ł	[NASQLITQRAQVSLLIRRELTERAKDFSLILDD
SKNPGYIKLRKIRAAQNISKTIATSQNRIYLTA DNLVINLQDESFTRGSDEIKGKK DNLVINLQDESFTRGSDEIKGKK SQFSLSQVLVDSABEGSLAAAAELAAQKREQ RLRKFRELHLMRNEARKLNHQEVVEEDKRL KLPANWEAKKARLEKEEKKKECAARG EDYEKVKLLEISAEDAERWERKKRKRNPDLG FSDYAAAQLRQYHRLTKQIKPDMETYEELBE KHGEEFFTSNSLLHGHTHVESTEEDBMYIDLE KQIEKROKYSRRRPYNDDADIDYNFRNAKF NKKAERFYGKYTAEIKQNLERGTAV OGGGTGSTGMRDSALTILLGIGPSHRNSLSIRL SQHSSPAMYSQTFHLUJG GSGRECNMAKTYDYLFKLLLIGDSGVGKTCVL FRFSEDAPMYSTGTHLUJG FSDYAAAQLRQYMGVATAEIKQNLERGTAV OGGGTGSTGMRDSALTILLGIGPSHRHSLSIRL SQHSSPAMYSQTFHLUJG GSGRECNMAKTYDYLFKLLLIGDSGVGKTCVL FRFSEDAPNSTFISTIGIDFKRITELDGKRIKLQ IWDTAGGRFRTITTAYYRGAMGIMUYVDIT NEKSFDNIRNWINNIEEHASADVEKMILGNKC DVNDKRQVSKERGEKLALDYGIKFMETSAK ANINVENAFFILARDIKAKMDKKLEGNSPQG SNQGVKITPDQQKRSSFFRCVLL PLLEAQPICANLVPYPTINATLDRITGKWFYI ASAFNRESYNKSVQCIAFFYYFTNKTEDTIF LREYQTRQDQCIYNTTYLNVQRENGTISRYV GGQEHFAHLLIRDTKTYMLAFDYNDEKNW GLSYYADKFYETKSCJAFFYTFDNKTEDTIF LREYQTRQDQCIYNTTYLNVQRENGTISRYV GGQEHFAHLLIRDTKTYMLAFDYNDEKNW GLSYYADKFYETKSCJAFFYTFDNKTEDTIF LREYQTRQDQCIYNTTYLNVQRENGTISRYV GGQEHFAHLLIRDTKTYMLAFDYNDEKNW GLSYADAVEFTIKEQIFFYTFDNKTEDTIF LREYQTRQDQCIYNTTYLNVQRENGTISRYV GGQEHFAHLLIRDTKTYMLAFDYNDEKNW GLSYADAVEFTIKEQIGFFYTEDVICEIPKSD VYYTDWKKDKCEPLEKQHEKERKQEEGES SVAFFPERVQLSQPONWNFSGAGGAWSLDF AGPADFFGIRLLSNFLMGDAVSDWSPMHEAA HGHQLSIRNLISQGWAVNITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQFESDLASV HEAARRGHVECVNSLTAGGNDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPESPLAYGQGHHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPESPLAJCHLERGEPPSIMQL CRLRIKCFGIQQHHKITKLVLPEDLKQFLH CRLRIKCFGIQQHHKITKLVLPEDLKQFLH CRLRIKCFGIQQHHKITKLVLPEDLKQFLH				!			VAITELSFSREYTAAVEAKQVAQQEAQRAQF
DNLVLNLQDESFTRGSDSLKGKK 982 SQFSLSQVLVDSAEGSLAAAELAAQKREQ RLRKFRELHLMRNBARKLNHQEVVEDKRL KLPANWFAKKARLEWELKEEKKKECAARG EDYEKVKLLEISAEAKERKKKECAARG EDYEKVKLLEISAEAKERKKKKCAARG EDYEKVKLLEISAEAKERKKKKCAPDLG FSDVAAAQLRQYHRLTKQIKPDMETYERLRE KQIEKROKYSRRRPYNDDADIDYNERNAKF NKKAERFYGKYTAEIKQNLERGTAV NKKAERFYGKYTAEIKQNLERGTAV VISTTVALITEKAPWILSAMTYTSVFLKFL QGGTGSTGMRDSALTLLGIGPSHRHSLSRL SQHSSPAPMYSQITHILVLG QGGTGSTGMRDSALTLLGIGPSHRHSLSRL SQHSSPAPMYSQITHILVLG SGRECINMAKTYDYLFKLLLIGDSGVGKTCVL FR'SEDAFNST'ISTIGIDFKIRTIELDGKRIKLQ IWDTAGQERFRTITTAYYRGAMGIMLVYDIT NEKSFDNIRNWIRNIEEHASADVEKMILGINKC DVNDKRQVSKERGEKLALDYGIKFMETSAK ANIVVENAFFTLARDKKMDKKLEGNSPQG SNQVKITPDQQKRSSFFRCVLL 937 2287 A 8039 393 311 EETHISENSYLEKYIPISANILTITIA PLLEAQIPLCANLVPYPITNATLDRITGKWFYI ASAFRNEEYNKSVQEIQATFFYPIPNKTEDTIF REYQTRQDQCIYNTTVLNVQRENGTISRYV GGQEHFAHLLIRDTKTYMLAFDYDDEKNW GLSYYADKPETTKEOLGEFYALDCLRIPKSD VYYTDWKKDKCEPLEKQHEKERKQEEGES SVYTDWKKDKCEPLEKQHEKERKQEEGES AGPRDFFGIRLLSNPLMGDAVSDWSPMHEAA INGIQLSVRADAVETVISLLORIPKSD VYYTDWKKDKCEPLEKQHEKERKQEEGES HEAARRGHVECVNSLIAVGGNDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKQQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELYPESPLAIVGLEREGFPSIMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQ	1						LVEKAKQEQRQKIVQAEGEAEAAKMLGEAL
DNLVLNLQDESFTRGSDSLKGKK 982 SQFSLSQVLVDSAEGSLAAAELAAQKREQ RLRKFRELHLMRNBARKLNHQEVVEDKRL KLPANWFAKKARLEWELKEEKKKECAARG EDYEKVKLLEISAEAKERKKKECAARG EDYEKVKLLEISAEAKERKKKKCAARG EDYEKVKLLEISAEAKERKKKKCAPDLG FSDVAAAQLRQYHRLTKQIKPDMETYERLRE KQIEKROKYSRRRPYNDDADIDYNERNAKF NKKAERFYGKYTAEIKQNLERGTAV NKKAERFYGKYTAEIKQNLERGTAV VISTTVALITEKAPWILSAMTYTSVFLKFL QGGTGSTGMRDSALTLLGIGPSHRHSLSRL SQHSSPAPMYSQITHILVLG QGGTGSTGMRDSALTLLGIGPSHRHSLSRL SQHSSPAPMYSQITHILVLG SGRECINMAKTYDYLFKLLLIGDSGVGKTCVL FR'SEDAFNST'ISTIGIDFKIRTIELDGKRIKLQ IWDTAGQERFRTITTAYYRGAMGIMLVYDIT NEKSFDNIRNWIRNIEEHASADVEKMILGINKC DVNDKRQVSKERGEKLALDYGIKFMETSAK ANIVVENAFFTLARDKKMDKKLEGNSPQG SNQVKITPDQQKRSSFFRCVLL 937 2287 A 8039 393 311 EETHISENSYLEKYIPISANILTITIA PLLEAQIPLCANLVPYPITNATLDRITGKWFYI ASAFRNEEYNKSVQEIQATFFYPIPNKTEDTIF REYQTRQDQCIYNTTVLNVQRENGTISRYV GGQEHFAHLLIRDTKTYMLAFDYDDEKNW GLSYYADKPETTKEOLGEFYALDCLRIPKSD VYYTDWKKDKCEPLEKQHEKERKQEEGES SVYTDWKKDKCEPLEKQHEKERKQEEGES AGPRDFFGIRLLSNPLMGDAVSDWSPMHEAA INGIQLSVRADAVETVISLLORIPKSD VYYTDWKKDKCEPLEKQHEKERKQEEGES HEAARRGHVECVNSLIAVGGNDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKQQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELYPESPLAIVGLEREGFPSIMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQ	ł		i	1			SKNPGYIKLRKIRAAQNISKTIATSQNRIYLTA
934 2284 A 8023 255 982 SQFSLSQVLVDSAEEGSLAAAAELAAQKREQ RLRKFRELHILMRNEARKLNHQEVVEEDKRI KLPANWEAKKARLEWELKEEEKKECAARG EDYEKVKLLEISAEDAERWERKKKECAARG EDYEKVKLLEISAEDAERWERKKKECAARG EDYEKVKLLEISAEDAERWERKKKECAARG EDYEKVKLLEISAEDAERWERKKKECAARG EDYEKVKLLEISAEDAERWERKKKECAARG EDYEKVKLLEISAEDAERWERKKKECAARG EDYEKVKLLEISAEDAERWERKKKECAARG EDYEKVKLLEISAEDAERWERKKKECAARG EDYEKVKLLEISAEDAERWERKKECAARG EDYEKVKLLEISAEDAERWERKKECAARG EDYEKVKLLEISAEDAERWERKKECAARG EDYEKVKLLEISAEDAERWERKKECAARG EDYEKVKLEISAEDAERWETKECH KOJEKOLE KOJE				:			
RIRKFRELHLMRNEARKLNHOEVVEEDKRL KLPANWEAKKARLEWELKEEEKKEKCAARG EDYEKVKLLEISAEDAERWERKKRKNPDLG FSDYAAAQLRQYHRLTKQIKPDMETYERLRE KHGEEFFPTSNSLLHGTHVPSTEEIDRMVIDLE KQIEKRDKYSRRPYNDDADIDYINERNAKF NKKAERFYGKYTAEIKQNLERGTAV NKKAERFYGKYTAEIKQNLERGTAV QCGGTGSTGMRDSALTLLGIGPSHRHSLSIRL SQHSSPAPMYSQTFHILVLG SGRECNMAKTYDYLFKLLLIGGDSVGKTCVL FRYSEDAFNSTFISTIGIDFKIRTIELDGKRIKLQ IWDTAGQERFRITITTAYYRGAMGIMLVYDIT NEKSFDIRRNWINNIEEHASADVEKMILGNKC DVNDKRQVSKERGEKLALDYGIKFMETSAK ANINVENAFFTLARDIKAKMDKKLEGNSPQG SNQGVKITPQQKRSFFRCVLL 937	934	2284	A	8023	255 .	982	
SLIPANWEAKKARLEWELKEEEKKKECAARG EDYEKVKLLEISAEDAERWERKKKKONPDLG FSDYAAAQLRQYHRLITKQIKPDMETTYERLRE KHGEBFFFTSNSLLHGTHYPSTEEIDRMVIDLE KQIEKRDKYSRRRPYNDDADIDYINERNAKF NKKAERFYGKYTAEIKQNLERGTAV NKKAERFYGKYTAEIKQNLERGTAV 935 2285 A 8027 59 310 LVSSTVILLITEKAPWISLAWTVTSYVFLKFL QGGGTGSTGMRDSALTLIGIGPSHRHSLSRIL SQHSSPAPMYSQTFHILVLG 936 2286 A 8032 1 639 SGRECIMAKTYDYLFKLLLIGDSGVGKTCVL FRFSEDAFNSTFISTIGIDFKIRTIELDGKRIKLQ IWDTAGQERFRITITATYRGAMGIMLVYDIT NEKSFDNIRNWIRNIEEHASADVEKMILGNKC DVNDKRQVSKERGEKLAUFGIKKEGNSPQG SNQGVKITPDQQKRSSFFRCVLL 937 2287 A 8039 393 311 EETHSENSYLEKYPISANLTITIA 938 2288 A 8052 675 1334 LHPAATSTAWLHVPPGLSMALSWVLTVLSLL PLLEAQIPLCANLVPVPITNATLDRITGKWFYI ASAFRNEEYNKSVQEIQATFFYFPNKTEDTIF LRYQTRQDQCIYNTYLNVQRENGISRYV GQQEHFAHLLIRDTKTYMLAFDVNDEKNW GLSVYADKPETTKEQLGEFYALDCLRIPKSD VYYTDWKCMCCEPLEKQHEKERKQEGES 939 2289 A 8055 12 1039 SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AGPRIPFGRILLSNGWAVNIITADHVSPLHEA CLGGHLSCVKLILKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLQHGASVQPSDLASS HEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKQQ DSPLHAVARTASEELACLLMDFGADTOAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLFEDLKQFLLH LL CRLRIRKCFGIQQHHKITKLVLFEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLFEDLKQF	1		-				
FSDYAAAQLRQYHRLTKQIKPDMETYERLRE KHGEEFFPTSNSLLHGTHVPSTEEIDRWIDLE KQIEKRDKYSRRRPYNDDADIDYNERNAKF NKKAERFYGKYTAEIKQNLERGTAV 935 2285 A 8027 59 310 LVSSTVNILTEKAPWNSLAWTVTSYVFLKFL QGGGTGSTGMRDSALTILGIGFSHRHSLSIRL SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFSHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQT	1		1				
FSDYAAAQLRQYHRLTKQIKPDMETYERLRE KHGEEFFPTSNSLLHGTHVPSTEEIDRWIDLE KQIEKRDKYSRRRPYNDDADIDYNERNAKF NKKAERFYGKYTAEIKQNLERGTAV 935 2285 A 8027 59 310 LVSSTVNILTEKAPWNSLAWTVTSYVFLKFL QGGGTGSTGMRDSALTILGIGFSHRHSLSIRL SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFSHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQTFHILVLG SQHSSPAPMYSQT	1		İ				(
State				1			1
SQIEKREKYSRRPYNDDADIDYINERNAKF NKKAEFYGKYTAEIKQNLERGTAV 935 2285 A 8027 59 310 LVSSTVALITEKAPWNSLAWTVTSYYFLKEL QGGTGSTGMRDSALTLLGIGPSHRHSLSIRL SQHSSFAPMYSQTFHILVLG SGRECNMAKTYDYLFKLLLIGGSGVGKTCVL FRFSEDAFNSTFISTIGIDFKIRTIELDGKRIKLQ IWDTAGQERFRTITTAYYRGAMGIMLVYDIT NEKSFDNIRNWINIEEHASADVEKMILGNKC DVNIDKRQVSKERGEKLALDYGIKFMETSAK ANINVENAFFTLARDIKAKMDKKLEGNSPQG SNQGVKITPDQQKRSSFFRCVLL 937 2287 A 8039 393 311 EETHISENSYLLEKYPISANLTLTIA 938 2288 A 8052 675 1334 LHPAATSTAWLHVPPGLSMALSWVLTVLSLL PLLEAQIPLCANLVPVITNATLDRITGKWFYI ASAFRNEEYNKSVQEIQATFYFTPNKTEDTIF LREYQTRQDQCIYNTTYLNVQRENGTISRYV GQEHFAHLLILRDTKTYMLAFDVNDEKNW GLSVYADKPTKEQI GEFVEALDCLRIPKSD VVYTDWKKDKCEPLEKQHEKERKQEEGES 939 2289 A 8055 12 1039 SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AGPRDFFGIRLSNPLMGDAVSDWSPMHEAA HIGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNILLQHGASVQPESDLASP HEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLH CRLRIRKCFGIQQHKTYKLTVLPEDLKQFLH CRLRIRKCFGIQQHKTYKLTVLPEDLKQFLH CRLRIRKCFGIQQHKTYKLTVLPEDLKQFLH CRLRIRKCFGIQQHKTYKLTVLPEDLKQFLH CRLRIRKCFGIQQHKTYKLTVLPEDLKQFLH CRLRIRKCFGIQQHKTYKLTVLPEDLKQFLH CRLRIRKCFGIQQHKTYKLTVLPEDLKQFLH CRLRIRKCFGIQQHKTYKL			ļ				
NKKAERFYGKYTAEIKQNLERGTAV							
935 2285 A 8027 59 310 LVSSTVNLLTEKAPWNSLAWTVTSYVFLKFL QGGGTGSTGMRDSALTLLGIGPSHRHSLSIRL SQHSSPAPMYSQTFHILVLG 936 2286 A 8032 1 639 SGRECNMAKTYDYLFKLLLIGDSGVGKTCVL FRYSEDAFNSTFISTIGIDFKIRTIELDGKRIKLQ IWDTAGQERFRITITATYRGAMGIMLVYDIT NEKSFDNIRNWIRNIEEHASADVEKMILGNKC DVNDKRQVSKERGEKLALDYGIKFMETSAK ANINVENAFFTLARDIKAKMDKKLEGNSPQG SNQCWITPDQQKRSSFFRCVLL 937 2287 A 8039 393 311 EETIHSENSYLLEKYIPISANLTITIA 938 2288 A 8052 675 1334 LHPAATSTAWLHVPPGLSMALSWVLTVLSLL PLLEAQIPLCANL VPVFITNATLDRITGKWFYI ASAFRNEEYNKSVQEIQATFFYFINKTEDTIF LREYQTRQDQCIVNITYLNVQRENGTISRYV GQEHFAHLLILRDTKTYMLAFDVNDEKNW GLSVYADKPETTKEQIGEFYEALDCLRIPKSD VYTDWKKDKCEPLEKQHEKERKQEEGES 939 2289 A 8055 12 1039 SSVAEFFERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGMDGSKP AGRDFFGIRLLSNPLMGDAVSDWSPMHEAA IHGHQLSLRNLISQQWAVNIITADHYSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVYSLLYGNDEHKISHLGT PLYLACEPOQRACVKKILLEGADVNQGKGQ DSPLHAVARTASEELACLLIMDFGADTQAKN AEGKRPVELVPPSSPLAQLFLERGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI							
936 2286 A 8032 1 639 SGRECNMAKTYDYLFKLLLIGDSGVGKTCVL FRFSEDAFNSTFISTIGIDFKIRTIELDGKRIKLQ IWDTAGQERFRITITAYYRGAMGIMLVYDIT NEKSFDNIRNWIRNIEEHASADVEKMILGKKC DVNDKRQVSKERGEKLALDYGIKFMETSAK ANINVENAFFILARDIKAKMDKKLEGNSPQG SNQGKITPDQQKRSSFFRCVLL 937 2287 A 8039 393 311 EETIHSENSYILEKYIPISANLTLTIA 938 2288 A 8052 675 1334 LHPAATSTAWLHVPPGLSMALSWVLTVLSLL PLLEAQIPLCANLVPVPITNATLDRITGKWFYI ASAFRNEEYNKSVQEIQATFFYFPNKTEDTIF LREYQTRQDQCIYNITTYLNVQRENGTISRYV GGQEHFAHLLILRDTKTYMLAFDVNDEKNW GLSYYADKPETTKEQLGFFYEALDCLRIPKSD VVYTDWKKDKCEPLEKQHEKERKQEEGES 939 2289 A 8055 12 1039 SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA IHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKQ DSSPLHAVARTASEELACLLIMDFGADTQAKN AEGKRPVELVPPESPLAQLFLERGPPSLMQL CRIRIKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI	935	2285	Α	8027	59	310	
SQHSSPAPMYSQTFHILVLG	}		-		7		
936 2286							
FRFSEDAFNSTFISTIGIDFKIRTIELDGKRIKLQ IWDTAGQERFRITITAYYRGAMGIMLYYDIT NEKSFDNIRNWIRNIEEHASADVEKMILGNKC DVNDKRQVSKERGEKLALDYGIKFMETSAK ANINVENAFFTLARDIKAKMDKKLEGNSPQG SNQGVKITPDQQKRSSFFRCVLL 937 2287 A 8039 393 311 EETHSENSYILEKYIPISANLTLTIA 938 2288 A 8052 675 1334 LHPA-ATSTAWLHVPPGLSMALSWVLTVLSLL PLLEAQIPLCANLVPVPITNATLDRITGKWFYI ASAFRNEEYNKSVQEIQAIFFYFTPNKTEDTIF LREYQTRQDQCIYNTTYLNVQRENGTISRYV GGVEHFAHLLILRDTKTYMLAFDVNDEKNW GLSVYADKPETTKEQLGEFYEALDCLRIPKSD VVYTDWKKDKCEPLEKQHEKERKQEEGES 939 2289 A 8055 12 1039 SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGGMDGSKP AGPRDPFGIRLLSNPLMGDAVSDWSPMHEAA HGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGYTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP HEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKQ DSPLHAVARTASEELACLLMDFGADTQAKN ABGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPSQPGGHLI	936	2286	A	8032	1	639	SCRECNMAKTYDVI EKI I I ICDSCVCKTOVI
IWDTAGQERFRTITTAYYRGAMGIMLVYDIT NEKSFDNIRNWIRNIEEHASADVEKMILGNKC DVNDKRQVSKERGEKLALDYGIKFMETSAK ANINVENAFFILARDIKAKMDKKLEGNSPQG SNQGVKITPDQQKRSSFFRCVLL 937 2287 A 8039 393 311 EETIHSENSYILEKYIPISANLTLTIA 938 2288 A 8052 675 1334 LHPAATSTAWLHVPPGLSMALSWVLTVLSLL PLLEAQIPLCANLVPVPITNATLDRITGKWFYI ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF LREYQTRQDQCIYNTTYLNVQRENGTISRYV GQQEHFAHLLILRDTKTYMLAFDVNDEKNW GLSVYADKPETTKEQLGEFYEALDCLRIPKSD VVYTDWKKDKCEPLEKQHEKERKQEEGES VVYTDWKKDKCEPLEKQHEKERKQEEGES VVYTDWKKDKCEPLEKQHEKERKQEEGES A 8055 12 1039 SSVAEFPPERVQLSQPQNWNFSGAGGAWSLDF AFQLKWSAELARLGESIMDGKQGGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA HIGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP HEAARRCHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVVLPEDLKQFLLH CRLRIRKCFGIQQHHKITKLVVLPEDLKQFLLH CRLRIRKCFGIQQHTMT CRLRIRKCFGIQQHTMT CRLRIRKCFGIQQHTMT CRLRIRCTH CRLRIRCTH CRLRIRCTH CRLRIRCTH CRLRIRCTH CRLRIRCTH CRLRIRCTH CRLRIRCTH CRLRIRCTH CRLRIRCTH CRLRIRCTH CRLRIRCTH CRLRIRCTH CRLRIRC	}			****	•		
937 2287 A 8039 393 311 EETIHSENSYILEKYIPISANLTLTIA 938 2288 A 8052 675 1334 LHPAATSTAWLHVPPGLSMALSWVLTVLSLL PLLEAQIPLCANLVPVPITINATLDRITGKWFYI ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF LREYQTRQDQCTYNTTYLNVQRENGTISRYV GQQEHFAHLLILRDTKTYMLAFDVNDEKNW GLSVYADKPETTKEQLGEFYEALDCLRIPKSD VVYTDWKKDKCEPLEKQHEKERKQEEGES 939 2289 A 8055 12 1039 SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA HGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP HEAAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI	[
DVNDKRQVSKERGEKLALDYGIKFMETSAK ANINVENAFFTLARDIKAKMDKKLEGNSPQG SNQGVKITPDQQKRSSFFRCVLL 937 2287 A 8039 393 311 EETIHSENSYILEKYIPISANLTLTIA 938 2288 A 8052 675 1334 LHPAATSTAWLHVPPGLSMALSWVLTVLSLL PLLEAQIPLCANLVPVPITNATLDRITGKWFYI ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF LREYQTRQDQCIYNTTYLNVQRENGTISRYV GQQEHFAHLLILRDTKTYMLAFDVNDEKNW GLSVYADKPETTKEQI.GEFYEALDCLRIPKSD VVYTDWKKDKCEPLEKQHEKERKQEEGES 939 2289 A 8055 12 1039 SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA IHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI	 						
ANINVENĀFFILARDIKAKMDKKLEGNSPQG SNQGVKITPDQQKRSSFFRCVLL 937 2287 A 8039 393 311 EETHSENSYILEKYIPISANLTLTIA 938 2288 A 8052 675 1334 LHPAATSTAWLHVPPGLSMALSWVLTVLSLL PLLEAQIPLCANLVPVPITNATLDRITGKWFYI ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF LREYQTRQDQCIYNITYLNVQRENGTISRYV GQQEHFAHLLILRDTKTYMLAFDVNDEKNW GLSVYADKPETTKEQI.GEFYEALDCLRIPKSD VVYTDWKKDKCEPLEKQHEKERKQEEGES 939 2289 A 8055 12 1039 SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA IHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI							
SNQGVKITPDQQKRSSFFRCVLL 937 2287 A 8039 393 311 EETIHSENSYILEKYIPISANLTLTIA 938 2288 A 8052 675 1334 LHPAATSTAWLHVPPGLSMALSWVLTVLSLL PLLEAQIPLCANLVPVPITNATLDRITGKWFYI ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF LREYQTRQDQCIYNTTYLNVQRENGTISRYV GGQEHFAHLLILRDTKTYMLAFDVNDEKNW GLSVYADKPETTKEQLGEFYEALDCLRIPKSD VVYTDWKKDKCEPLEKQHEKERKQEEGES 939 2289 A 8055 12 1039 SSVAEFPERVQLSOPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA IHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI							
937 2287 A 8039 393 311 EETHSENSYILEKYIPISANLTLTIA 938 2288 A 8052 675 1334 LHPAATSTAWLHVPPGLSMALSWVLTVLSLL PLLEAQIPLCANLVPVPITINATLDRITGKWFYI ASAFRNEEYNKSVQEIQAFFYFTPNKTEDTIF LREYQTRQDQCIYNTTYLNVQRENGTISRYV GQQEHFAHLLILRDTKTYMLAFDVNDEKNW GLSVYADKPETTKEQI.GEFYEALDCLRIPKSD VVYTDWKKDKCEPLEKQHEKERKQEEGES 939 2289 A 8055 12 1039 SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA IHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI	Į I			1			
938 2288 A 8052 675 1334 LHPAATSTAWLHVPPGLSMALSWVLTVLSLL PLLEAQIPLCANLVPVPITNATLDRITGKWFYI ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF LREYQTRQDQCIYNTTYLNVQRENGTISRYV GQEHFAHLLILRDTKTYMLAFDVNDEKNW GLSVYADKPETTKEQI.GEFYEALDCLRIPKSD VVYTDWKKDKCEPLEKQHEKERKQEEGES VYYTDWKKDKCEPLEKQHEKERKQEEGES SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA IHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNILLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L	037	2287	Α	9020	202	211	
PLLEAQIPLCANLVPVPITNATLDRITIGKWFYI ASAFRNEEYNKSVQEIQAIFFYFTPNKTEDTIF LREYQTRQDQCIYNTTYLNVQRENGTISRYV GQQEHFAHLLILRDTKTYMLAFDVNDEKNW GLSVYADKPETTKEQI.GEFYEALDCLRIPKSD VVYTDWKKDKCEPLEKQHEKERKQEEGES 939 2289 A 8055 12 1039 SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA IHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI							
ASAFRNEEYNKSVQEIQATFFYFTPNKTEDTIF LREYQTRQDQCIYNTTYLNVQRENGTISRYV GQQEHFAHLLILRDTKTYMLAFDVNDEKNW GLSVYADKPETTKEQI.GEFYEALDCLRIPKSD VVYTDWKKDKCEPLEKQHEKERKQEEGES 939 2289 A 8055 12 1039 SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA IHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHI.SCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI	220	2400	^	0032	0/3	1334	
P40 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI]						
GGQEHFAHLLILRDTKTYMLAFDVNDEKNW GLSVYADKPETTKEQI.GEFYEALDCLRIPKSD VVYTDWKKDKCEPLEKQHEKERKQEEGES 939 2289 A 8055 12 1039 SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA IHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI							
GLSVYADKPETTKEQI.GEFYEALDCLRIPKSD VVYTDWKKDKCEPLEKQHEKERKQEEGES 939 2289 A 8055 12 1039 SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA IHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L							
939 2289 A 8055 12 1039 SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA IIHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI	ļ						
939 2289 A 8055 12 1039 SSVAEFPERVQLSQPQNWNFSGAGGAWSLDF AEQLKWSAELARLGESIMDGKQGGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA IIHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI							
AEQLKWSAELARLGESIMDGKQGGMDGSKP AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA IHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHISCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI	020	-0000		00.55			VVYIDWKKDKCEPLEKQHEKERKQEEGES
AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA IHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI	939	2289	A	8055	12	1039	
IHGHQLSLRNLISQGWAVNIITADHVSPLHEA CLGGHISCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI							
CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI							AGPRDFPGIRLLSNPLMGDAVSDWSPMHEAA
CLGGHLSCVKILLKHGAQVNGVTADWHTPL FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI							IHGHQLSLRNLISQGWAVNIITADHVSPLHEA
FNACVSGSWDCVNLLLQHGASVQPESDLASP IHEAARRGHVECVNSLIAYGGNIDHKISHLGT PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI							
iheaarrghvecvnsliayggnidhkishlgt Plylacenqqracvkkllesgadvnqgkgq Dsplhavartaseelacllmdfgadtqakn Aegkrpvelvppesplaqlfleregppslmql Crlrirkcfgiqqhhkitklvlpedlkqfllh L 940 2290 A 8058 2 1203 KVLSirepahstarkasepsqpsqpsqpgghli					ļ		
PLYLACENQQRACVKKLLESGADVNQGKGQ DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI		•	. [1		
DSPLHAVARTASEELACLLMDFGADTQAKN AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI		ļ			l		
AEGKRPVELVPPESPLAQLFLEREGPPSLMQL CRLRIRKCFGIQQHHKITKLVLPEDLKQFLLH L 940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI		[l		
940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI			ļ	}	ļ		
940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI				 			
940 2290 A 8058 2 1203 KVLSIREPAHSTARKASEPSQPSQPSQPGGHLI				1	i		
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	940	2290	A	8058	2	1203	
	' ',				-	.203	
	LI						ALCENTALIER I SELGILISDIS MECHSSICI

SEQ ID	SEQ ID	Met	1 650	Predicted	I 70 - 21 - 4 - 4 - 4	
NO: of	NO: of	hod	SEQ ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid.
nucl-	peptide		in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	İ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ľ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ļ	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		i		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ļ	l	J	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			1	peptide	i •	/=possible nucleotide deletion, \=possible
Ĺ	1			sequence		nucleotide insertion
						VVDTVMCPNMPNKSVLLYTLSFIYIFIFVIGMI
ł	i	ł	Į.	İ		ANSVVVWVNIQAKTTGYDTHCYILNLALADL
			ĺ	1		WVVLTIPVWVVSLVQHNQWPMGELTCKVTH
						LIFSINLFGSIFFLTCMSVDRYLSITYFTNTPSS
		!			i	RKKMVRRVVCILVWLLAFCVSLPDTYYLKT
	İ	1	{		f	VTSASNNETYCRSFYPEHSIKEWLIGMELVSV
	1			İ		VLGFAVPFSIIAVFYFLLARAISASSDQEKHSS
						RKIIFSYVVVFLVCWLPYHVAVLLDIFSILHYI
	·	l	1	İ		PFTCRLEHALFTALHVTQCLSLVHCCVNPVL
	1					YSFINRNYRYELMKAFIFKYSAKTGLTKLIDA
941	2291	A -	8059	73	433	SRVSETEYSALEQSTK
741	2271	^	8039	/3	432	DMAGLMTIVTSLLFLGVCAHHIIPTGSVVLPS
		ĺ				PCCMFFVSKRIPENRVVSYQLSSRSTCLKAGV
						IFTTKKGQQFCGDPKQEWVQRYMKNLDAKQ KKASPRARAVAVKGPVQRYPGNQTTC
942	2292	A	8067	278	1262	GGIGEIKQRPSCLGRCLDPSLSVLMNISLGLGS
7,2	2272	Λ.	0007	270	1202	VFSAVISQKPSRDICQRGTSLTIQCQVDSQVT
						MMFWYRQQPGQSLTLIATANQGSEATYESGF
						VIDKFPISRPNLTFSTLTVSNMSPEDSSIYLCSA
			i i			GRQGTYEQYFGPGTRLTVTEDLKNVFPPEVA
						VFEPSEAEISHTQKATLVCLATGFYPDHVELS
						WWVNGKEVHSGVSTDPOPLKEOPALNDSRY
						CLSSRLRVSATFWQNPRNHFRCQVQFYGLSE
						NDEWTQDRAKPVTQIVSAEAWGRADCGFTS
				•		ESYQQGVLSATILYEILLGKATLYAVLVSALV
						LMAMVKRKDSRG
943	2293	Α	8070	1	879	MVKVVPATRGNLPRSQLTGTHQHCQPREPKI
						TASERLRRRPRATARLRAHAAPPEPPLAVFAP
				•		PSDRKELLALPVACDPVIASVMSWVQAASLI
	}					QGPGDKGDVFDEEADESLLAQREWQSNMQR
						RVKEGYRDGIDAGKAVTLQQGFNQGYKKGA
						EVILNYGRLRGTLSALLSWCHLHNNNSTLINK
						INNLLDAVGQCEEYVLKHLKSITPPSHVVDLL
1	[ĺ		i		DSIEDMDLCHVVPAEKKIDEAKDERLCENNA
	ı	i				EFNKNCSKSHSGIDCSYVECCRTQEHAHSGK
944	2294	A -	8073		797	PKPHMDFGTDSQF
· · ·		^	3073	* 1	191	ESARWSRQLRRTLIRLSFPISCGRSHAFGGCK MAATSGTDEPVSGELVSVAHALSLPAESYGN
		i				DPDIEMAWAMRAMQHAEVYYKLISSVDPOF
İ			1			LKLTKVDDQIYSEFRKNFETLRIDVLDPEELK
!			}		•	SESAKEKWRPFCLKFNGIVEDFNYGTLLRLD
1		ľ	l			CSQGYTEENTIFAPRIOFFAIEIARNREGYNKA
ŀ				1		VYISVQDKEGEKGVNNGGEKRADSGEEENT
İ		1				KNGGEKGADSGEEKEEGINREDKTDKGGEK
ľ	ľ	1	· '		1	GKEADKEINKSGEKAM
945	2295	A	8074	2	505	GAATLLRSASSAARKAAEAEQVWLHLHRYL
						SADRRVLGLREWGRPASERECSLCORLKREL
}	-				1	NMGDVEKGKKIFIMKCSOCHTVEKGGKHKT
l	}		.	i	j	GPNLIIGLFGRKTGQAPGYSYTAANKNKGIIW
	İ	İ				GEDTLMEYLENPKKYIPGTKMIFVGIKKKEER
		•			ŧ	ADLIAYLKKATNE
946	2296	A	8081	42	590	EGRRGKFGGKLCNFLFYFHSNSAESRMDVLF
	ł					VAIFAVPLILGQEYEDEERLGEDEYYQVVYY
]		i	ł	YTVTPSYDDFSADFTIDYSIFESEDRLNRLDK
	[1			i	DITEAIETTISLETARADHPKPVTVKPVTTEPQ
		1			j	SPRSEAMPCPVLRSPIPLPPVRVPLFRWGCISC
	[İ		ĺ	ł	KKVGRRLLMTLWMGVWQEEIGR
947	2297	A	8084	322	549	GGGSSPRELAGAAGLTVTSQAVAARRQQPSF
				ł	· · · · · · · · · · · · · · · · · · ·	SRARAPAHSLRAALSLASSARSWGAVSRDRG
	·					The state of the s

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
948	2298	В	8093	3905	846	PCPPAIMYQSSNKC MEPGEVKDRILENISLSVKKLQSYFAACEDEI PAIRNHDKVLQRLCEHLDHALLYGLQDLSSG YWVLVVHFTREAIKQIEVLQHVATNLGRSR AWLYLALNENSLESYLRLFQENLGLLHKYYV KNALVCSHDHLTLFLTLVSGLEFIRFELDLDA PYLDLAPYMPDYYKPQYLLDFEDRLPSSVHG SDSLSLNSFNSVTSTNLEWDDSAIAPSSEDYD FGDVFPAVPSVPSTDWEDGDLTDTVSGPRST ASDLTSSKASTRSPTQRQNPFNEPAETVSSS DTTPVHTTSQEKEEAQALDPPDACTELEVIRV TKKKKIGKKKKSRSDEEASPLHPACSQKKCA KQGDGDSRNGSPSLGRDSPDTMLASPQEEGE GPSSTTESSERSEPGLLIPEMKDTSMERLGQPL SKVIDQLNGQLDPSTWCSRAEPPDQSFRTGSP GDAPERPPLCDFSEGLSAPMDFYRFTVESPST VTSGGGHHDPAGLGQPLHVPSSPEAAGQEEE GGGGEGQTPRPLEDTTREAQELEAQLSLVRE GPVSEPEPGTQEVLCQLKRDQPSPCLSSAEDS GVDEGQGSPSEMYHSSEFRVDNNHLLLMIH VFRENEEQLFKMIRMSTGHMEGNLQLLYVLL TDCYVYLLRKGATEKPYLVEEAVSYNELDY VSVGLDQQTVKLVCTNRRKOFLLDTADVAL AEFFLASLKSAMIKGCREPPYPSILTDATMEK LALAKFVAQESKCEASAVTVRFYGLVHWED PTDESLGPTPCHCSPPEGTITKEGMLHYKAGT SYLGKEHWKTCFVVLSNGILYQYPDRTDVIP LLSVNMGGEQCGGCRRANTTDRPHAFQVILS DPPCLELSAESEAEMAEWMQHLCQAVSKGVI PQGVAPSPCIPCCLVLTDDRLFTCHEDCQTSF FRSLGTAKLGDISAVSTEPGKEYCVLEFSQDS QQLLPPWVIYLSCTSELDRLLSALNSGWKTIY QVDLPPHTAIQEASNKKKFEDALSLIHSAWQR SDSLCRGRASRDPWC*
949	2299	A	8095	9	2374	ARRADTVLLESPSMLQGLLPVSLLLSVAVSAI KELPGVKKYEVVYPIRLHPLHKREAKEPEQQ EQFETELKYKMTINGKIAVLYLKKNKNLLAP GYTETYYNSTGKEITTSPQIMDDCYYQGHILN EKVSDASISTCRGLRGYPSQGDQRYFIEPLSPI HRDGQEHALFKYNPDEKNYDSTCGMDGVL WAHDLQQNIALPATKLVKLKDRKVQEHEKY IEYYLVLDNGEFKRYNENQDEIRKRVFEMAN YVNMLYKKLNTHVALVGMEIWTDKDKIKIT PNASFTLENFSKWRGSVLSRRKHDIAQLITA TELAGTTVGLAFMSTMCSPYSVGVVQDHSD NLLRVAGTMAHEMGHNFGMFHDDYSCKCPS TICVMDKALSFYIPTDFSSCSRLSYDKFFEDKL SNCLFNAPLPTDIISTPICGNQLVEMGEDCDC GTSEECTNICCDAKTCKIKATFQCALGECCEK CQFKKAGMVCRPAKDECDLPEMCNGKSGNC PDDRFQVNGFPCHHGKGHCLMGTCPTLQEQ CTELWGPGTEVADKSCYNRNEGGSKYGYCR RVDDTLIPCKANDTMCGKLFCQGGSDNLPW KGRIVTFLTCKTFDPEDTSQEIGMVANGTKCG DNKVCINAECVDIEKAYKSTNCSSKCKGHAV CDHELQCQCEEGWIPPDCDDSSVVFHFSIVVG VLFPMAVIFVVVAMVIRHQSSREKQKKDQRP LSTTGTRPHKQKRKPQMVKAVQPQEMSQMK PHYYDLPVEGNPPASFHKDTNALPPTVFKD NPMSTPKDSNPKA

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide election, \=possible nucleotide insertion
950	2300	A	8100	1	1251	MGLLLMILASAVLGSFLTLLAQFFLLYRRQPE PPADEAARAGEGFRYIKPVPGLLLREYLYGG GRDEPSGAAPEGGATPTAAPETPAPPTRETC YFLNATILFLFRELRDTALTRRWVTKKIKVEF EELLQTKTAGRLLEGLSLRDVFLGETVPFIKTI RLVRPVVPSATGEPDGPEGEALPAACPEELAF EAEVEYNGGFHLAIDVDLVFGKSAYLFVKLS RVVGRLRLVFTRVPFTHWFFSFVEDPLIDFEV RSQFEGRPMPQLTSIIVNQLKKIIKRKHTLPNY KIRFKPFFPYQTLQGFEEDEEHIHIQQWALTE GRLKVTLLECSRLLIFGSYDREANVHCTLELS SSVWEEKQRSSIKTGTISLTAVFMGWHRVSE AFPGLWYKLLVDLPFWGLEDGGPLLTVPLRQ CPG
951	2301	A	8108	1612	839	EVALFCFEMAAGMYLEHYLDSIENLPFELQR NFQLMRDLDQRTEDLKAEIDKLATEYMSSAR SLSSEEKLALLKQIQEAYGKCKEFGDDKVQL AMQTYEMVDKHIRRLDTDLARFEADLKEKQI ESSDYDSSSSKGKKKGRTQKEKKAARARSKG KNSDEEAPKTAQKKLKLVRTSPEYGMPSVTF GSVHPSDVLDMPVDPNEPTYCLCHQVSYGE MIGCDNPDCSIEWFHFACVGLTTKPRGKWFC PRCSQERKKK
952	2302	A	8112	595	291	PSVASIARRFSGRALWPPSHSVPGNRALCPRL LHGTTLPGGNQRELARQKNMKKQSDSVKGK RRDDGLSAAARKQRDSTPRDSEIMQQKQKK ANEKKEEPK
953	2303	A	8118	1	669	VCAGIRDPCSTPLAKPAAGGAENLSFGKQPG LETNILKMTTPNKTPPGADPKQLERTGTVREI GSQAVWSLSSCKPGFGVDQLRDDNLETYWQ SDGSQPHLVNIQFRRKTTVKTLCIYADYKSDE SYTPSKISVRVGNNFHNLQEIRQLELVEPSGW IHVPLTDNHKKPTRTFMIQIAVLANHQNGRD THMRQIKIYTPVEESSIGKFPRCTTIDFMMYRS IR
954	2304	A	8133	66	1015	PPLPPRSFPNLFSRPEPLPEPGRRGCNRSREPA ARAPSPPPFEGAPGRAMVKVTFNSALAQKE AKKDEPKSGEEALIIPPDAVAVDCKDPDDVV PVGQRRAWCWCMCFGLAFMLAGVILGGAY LYKYFALQPDDVYYCGIKYIKDDVILNEPSAD APAALYQTIEENIKIFEEEEVEFISVPVPEFADS DPANIVHDFNKKLTAYLDLNLDKCYVIPLNT SIVMPPRNLLELLNIKAGTYLPQSYLIHEHMV ITDRIENIDHLGFFIYRLCHDKETYKLQRRETI KGIQKREASNCFAIRHFENKFAVETLICS
955	2305	A	8143	1854	798	VESRSAWHEGEDQIDRLDFIRNQMNLI.TLDV KKKIKEVTEEVANKVSCAMTDEICRLSVLVD EFCSEFHPNPDVLKIYKSELNKHIEDGMGRNL ADRCTDEVNALVLQTQQEIIENLKPLLPAGIQ DKLHTLIPCKKFDLSYNLNYHKLCSDFQEDIV FRFSLGWSSLVHRFLGPRNAQRVLLGLSEPIF QLPRSLASTPTAPTTPATPDNASQEELMITLVT GLASVTSRTSMGIIIVGGVIWKTIGWKLLSVS LTMYGALYLYERLSWTTHAKERAFKQQFVN YATEKLRMIVSSTSANCSHQVKQQIATTFARL CQQVDITQKQLEEEIARLPKEIDQLEKIQNNS KLLRNKAVQLENELENFTKQFLPSSNEES ASGSPAPSSSSAMAAACGPGAAGYCLLIGLH
730	2300	^	1,710	1034	170	ASGSPAPSSSSAMAAACGPGAAGYCLLLGLH LFLLTAGPALGWNDPDRMLLRDVKALTLHY

SEQ ID	SEQ ID	Met	SEQ	Deadistad	Deadisted and	A in cold and the first of the cold and the
NO: of	NO: of	hod	ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ľ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	Í	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
l	ł	ł	Į	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
}			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
}			ł	peptide	·	/=possible nucleotide deletion, \=possible
	L	<u> </u>	<u> </u>	sequence	l_	nucleotide insertion
						DRYTTSRRLDPIPQLKCVGGTAGCDSYTPKVI
1					Ì	QCQNKGWDGYDVQWECKTDLDIAYKFGKT
ł	ì	ł	l			VVSCEGYESSEDQYVLRGSCGLEYNLDYTEL
ļ	1					GLQKLKESGKQHGFASFSDYYYKWSSADSC
Ì	-					NMSGLITIVVLLGIAFVVYKLFLSDGQYSPPP
i	ł	l		}		YSEYPPFSHRYQRFTNSAGPPPPGFKSEFTGPQ
	ŀ	ł				NTGHGATSGFGSAFTGQQGYENSGPGFWTGL
l			[GTGGILGYLFGSNRAATPFSDSWYYPSYPPSY PGTWNRAYSPLHGGSGSYSVCSNSDTKTRTA
			1			SGYGGTRRR
957	2307	A	8159	1492	528	THVVMTGMCYAPHQVLSYINGVTTSKPGVSL
		١.,	0.37	1472	320	VYSMPSRNLSLRLEGLQEKDSGPYSCSVNVO
						DKQGKSRGHSIKTLELNVLVPPAPPSCRLQGV
]]	i i			PHVGANVTLSCQSPRSKPAVQYQWDRQLPSF
			1			QTFFAPALDVIRGSLSLTNLSSSMAGVYVCKA
	İ				!	HNEVGTAQCNVTLEVSTGPGAAVVAGAVVG
İ]]			TLVGLGLLAGLVLLYHRRGKALEEPANDIKE
			[]			DAIAPRTLPWPKSSDTISKNGTLSSVTSARAL
ł		İ	1			RPPHGPPRPGALTPTPSLSSQALPSPRLPTTDG
						AHPQPISPIPGGVSSSGLSRMGAVPVMVPAQS
						QAGSLV
958	2308	Α	8161	2340	1192	ELARRPKQQSSEKSRNMIRNWLTIFILFPLKLV
				-	·	EKCESSVSLTVPPVVKLENGSSTNVSLTLRPP
						LNATLVITFEITFRSKNITILELPDEVVVPPGVT
						NSSFQVTSQNVGQLTVYLHGNHSNQTGPRIR
				.		FLVIRSSAISIINQVIGWIYFVAWSISFYPQVIM
					1	NWRRKSVIGLSFDFVALNLTGFVAYSVFNIGL
						LWVPYIKEQFLLKYPNGVNPVNSNDVFFSLH
				_		AVVLTLIIIVQCCLYERGGQRVSWPAIGFLVL
					ľ	AWLFAFVTMIVAAVGVITWLQFLFCFSYIKL AVTLVKYFPQAYMNFYYKSTEGWSIGNVLL
			İ			DFTGGSFSLLQMFLQSYNNDQWTLIFGDPTK
			1		i	FGLGVFSIVFDVVFFIQHFCLYRKRPGYDQLN
959	2309	A	8163	521	1345	GERAGRRGRLGVWAQPQPLLPRPVGSRRE
					10.0	MQPPGPPPAYAPTNGDFTFVSSADAEDLSGSI
						ASPDVKLNLGGDFIKESTATTFLRQRGYGWL
		l	Ī	İ		LEVEDDDPEDNKPLLEELDIDLKDIYYKIRCV
						LMPMPSLGFNRQVVRDNPDFWGPLAVVLFFS
		J		. [MISLYGQFRVVSWIITIWIFGSLTIFLLARVLG
		1	1	l	[GEVAYGQVLGVIGYSLLPLIVIAPVLLVVGSF
		1		- 1	l	EVVSTLIKLFGVFWAAYSAASLLVGEEFKTK
960	2210	_	9165			KPLLIYPIFLLYIYFLSLYTGV
300	2310	A	8167	1	2921	MTCFKGQKGEQRSHAFEANKDHKAKVPSPN
.	1	- 1				LYSQLNALQFTVDERSILWLNQFLLDLKQSL
	i	I		ł	Ì	NOFMAVYKLNDNSKSDEHVDVRVDGLMLK
		ì		1	l	FVIPSEVKSECHQDQPRAISIQSSEMIATNTRH
	l	1	1	1	1	CPNCRHSDLEALFQDFKDCDFFSKTYTSFPKS
1	Į			į		CDNFNLLHPIFQRHAHEQDTKMHEIYKGNITP
1	Ī		l		l l	QLNKNTLKTSAATDVWAVYFSQFWIDYEGM KSCKCDDISEVDSEDI SIMICODTRYAESOVED
ł	ł	ł		ļ	1	KSGKGRPISFVDSFPLSIWICQPTRYAESQKEP QTCNQVSLNTSQSESSDLAGRLKRKKLLKEY
	}	I			l	YSTESEPLTNGGQKPSSSDTFFRFSPSSSEADI
l	ł	ĺ	1	ľ	1	HLLVHVHKHVSMQINHYQYLLLLFLHESLILL
ļ		i	1		ļ	SENLRKDVEAVTGSPASQTSICIGILLRSAELA
ļ	ł	l	ŀ		ļ	LLLHPVDQANTLKSPVSESVSPVVPDYLPTEN
i	ĺ	ĺ	1	1	ì	GDFLSSKRKQISRDINRIRSVTVNHMSDNRSM
ŀ	j	1		1	l	SVDLSHIPLKDPLLFKSASDTNLOKGISFMDY
l	l	- 1	ł		l	LSDKHLGKISEDESSGLVYKSGSGEIGSETSD
!	. !	i		İ	ľ	KKDSFYTDSSSVLNYREDSNILSFDSDGNQNI
	'	I		İ		LSSTLTSKGNETIESIFKAFDLLPEAASLSENL

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion DISKEETPPVRTLKSQSSLSGKPKERCPPNLAP LCVSYKNMKRSSSQMSLDTISLDSMILEEQLL ESDGSDSHMFLEKGNKKNSTTNYRGTAESVN AGANLQNYGETSPDAISTNSEGAQENHDDLM SVVYFKITGVNGEIDIRGEDTEICLQVNQVTP DQLGNISLRHYLCNRPVGSDQKAVIHSKSSPE ISLRFESGPGAVIHSLLAEKNGFLQCHIENFST
						EFL'ISSLMNIQHFLEDETVATVMPMKIQVSNT KINLKDDSPRSSTVSLEPAPVTV:HIDHLVVER SDDGSFHIRDSHMLN:TGNDLKENVKSDSVLL TSGKYDLKKQRSVTQATQTSPGVPWPSQSAN FPEFSFDFTREQLMEENESLKQELAKAKMAL AEAHLEKDALLHHIKKMTVE
961	2311	A	8172	1442	682	TAAMSIFTPTNQIRLTNVAVVRMKRAGKRFEI ACYKNKVVGWRSGVEKDLDEVLQTHSVFVN VSKGQVAKKEDLISAFGTDDQTEICKQILTKG EVQVSDKERHTQLEQMFRDIATIVADKCVNP ETKRPYTVILIERAMKDIHYSVKTNKSTKQQA LEVIKQLKEKMKIERAHMRLRFILPVNEGKKL KEKLKPLIKVIESEDYGQQLEIVCLIDPGCFREI DELIKKETKGKGSLEVLNLKDVEEGDEKFE
962	2312	A	8175	286	587	NISNKAEVSSHPSVISHSMDSFGQPRPEDNQS VLRRMQKKYWKTKQVFIKATGKKEDEHLVA SDAELDAKLEVFHSVQETCTELLKIIEKYQLR LNGMKS
963	2313	Ą	8181	13	2215	AEGCAERRGTEPVVELSMSWESGAGPGLGSQ GMDLVWSAWYGKCVKGKGSLPLSAHGIVV AWLSRAEWDQVTVYLFCDDHKLQRYALNRI TVWRSRSGNELPLAVASTADLIRCKLLDVTG GLGTDELRLLYGMALVRFVNLISERKTKFAK VPLKCLAQEVNIPDWIVDLRHELTHKKMPHI NDCRRGCYFVLDWLQKTYWCRQLENSLRET WELEEFREGIEEEDQEEDKNIVVDDITTEQKPE PQDDGKSTESDVKADGDSKGSEEVDSHCKK ALSHKELYERARELLVSYEEEQFTVLEKFRYL PKAIKAWNNPSPRVECVLAELKGVTCENREA VLDAFLDDGFLVPTFEQLAALQEYEENVDL NDVLVPKPFSQFWQPLLRGLHSQNFTQALLE RMLSELPALGISGIRPTYILRWTVELIVANTKT GRNARRFSAGQWEARRGWRLFNCSASLDWP RMVESCLGSPCWASPQLLRIIFKAMGQGLPD EEQEKLLRICSIYTQSGENSLVQEGSEASPIGK SPYTLDSLYWSVKPASSSFGSEAKAQQQEEQ GSVNDVKEEEKEEKEVLPDQVEEEEENDDQE
	٠					EEEEDEDDEDDEEEDRMEVGPFSTGQESPTA ENARLLAQKRGALQGSAWQVSSEDVRWDTF PLGRMPGQTEDPAELMLENYDTMYLLDQPV LEQRLEPSTCKTDTLGLSCGVGSGNCSNSSSS NFEGLLWSQGQLHGLKTGLQLF
964	2314	A	8184	6	1393	EPRRNI'RDDSTRPRTRGRTRGRRRRACRSAE GTGLRSLLLPPRLQLPAGPFSRCRWDPVSSPR PSTMPPKKGGDGIKPPPIIGRFGTSLKIGIVGLP NVGKSTFFNVLTNSQASAENFPFCTIDPNESR VPVPDERFDFLCQYHKPASKIPAFLNVVDIAG LVKGAHNGQGLGNAFLSHISACDGIFHLTRA FEDDDITHVEGSVDPIRDIEIHEELQLKDEEMI GPIIDKLEKVAVRGGDKKLKPEYDIMCKVKS WVIDQKKPVRFYHDWNDKEIEVLNKHLFLTS KPMVYLVNLSEKDYIRKKNKWLIKIKEWVD KYDPGALVIPFSGALELKLQELSAEERQKYLE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion ANMTQSALPKIIKAGFAALQLEYFFTAGPDEV RAWTIRKGTKAPQAAGKIHTDFEKGFIMAEV MKYEDFKEEGSENAVKAAGKYRQQGRNYIV EDGDIIFFKFNTPQQPKKK RSFSLSFSLLSPSEMMALGAAGATRVFVAMV AAALGGHPLLGVSATLNSVLNSNAIKNLPPPL GGAAGHPGSAVSAAPGILYPGGNKYQTIDNY QPYPCAEDBECGTDEYCASPTRGGDAGVQIC
						LACRKRRKRCMRHAMCCPGNYCKNGICVSS DQNHFRGEIEETITESFGNDHSTLDGYSRRTT LSSKMYHTKGQEGSVCLRSSDCASGLCCARH FWSKICKPVLKEGQVCTKHRRKGSHGLEIFQ RCYCGEGI SCRIOK DHHOA SNSSRI HTCORH
966	2316	A	8207	416	4082	RCYCGEGLSCRIQKDHHQASNSSRLHTCQRH KFKLIKIMLLTLIILLPVVSKFSFVSLSAPQHW SCPEGTLAGNGNSTCVGPAPFLIFSHGNSIFRI DTEGTNYEQLVVDAGVSVIMDFHYNEKRIY WVDLERQLLQRVFLNGSRQERVCNIEKNVSG MAINWINEEVIWSNQQEGIITVTDMKGNNSHI LLSALKYPANVAVDPVERFIFWSSEVAGSLY RADLDGVGVKALLETSEKITAVSLDVLDKRL FWIQYNREGSNSLICSCDYDGGSVHISKHPTQ HNLFAMSLFGDRIFYSTWKMKTIWIANKHTG KDMVRINLHSSFVPLGELKVVHPLAQPKAED DTWEPEQKLCKLRKGNCSSTVCGQDLQSHLC MCAEGYALSRDRKYCEGNDWKYCEDVNEC AFWNHGCTLGCKNTPGSYYCTCPVGFVLLPD GKRCHQLVSCPRNVSECSHDCVLTSEGPLCF CPEGSVLERDGKTCSGCSSPDNGGCSQLCVPL SPVSWECDCFPGYDLQLDEKSCAASGPOPFL LFANSQDIRHMHFDGTDYGTLLSQQMGMVY ALDHDPVENKTYFAHTALKWIERANMDGSQ RERLIEEGVDVPEGLAVDWIGRRFYWTDRGK SLIGRSDLNGKRSKIITIENISQPRGIAVHPMAK RLFWTDTGINPRIESSSLQGLGRLVIASSDLIW PSGITIDFLTDKLYWCDAKQSVIEMANLDGSK RRRLTQNDVGHPFAVAVFEDYWFSDWAMP SVIRVNKRTGKDRVRLQGSMLKPSSLVVVHP LAKPGADPCLYQNGGCEHICKKRLGTAWCS CREGFMKASDGKTCLALDGHQLLAGGEVDL KNQVTPLDILSKTRVSEDNITESQHMLVAEIM VSDQDDCAPVGCSMYARCISEGEDATCQCLK GFAGDGKLCSDIDECEMGVPVCPPASSKCINT EGGYVCRCSEGYQGDGIHCLDIDECQLGVHS CGENASCTNTEGGYTCMCAGRLSEPGLICPD STPPPHILREDDHIIYSVRNSDSECPLSHDGYCL HDGVCMYIEALDKYACNCVVGYIGERCQYR DLKWWELRHAGHGQQQKVIVVAVCVVVLV MLLLLSLWGAHYYRTQKLLSKNPKNPYEESS RDVRSRRPADTEDGMSSCPQPWFVVIKEHQD LKNGGQPVAGEDGQAADGSMQPTSWRQEPQ LCGMGTEQGCWIPVSDKGSCPQVMERSFH MPSYGTQTLEGGVEKPHSLLSANPLWQQRAL DPPHQMELTQ
967	2317	A	8210	3	601	SSAMGSRSSHAAVIPDGDSIRRETGFSQASLL RLHHRFRALDRNKKGYLSRMDLQQIGALAV NPLGDRIIESFFPDGSQRVDFPGFVRVLAHFRP VEDEDTETQDPKKPEPLNSRRNKLHYAFQLY DLDRDGKISRHEMLQVLRLMVGVQVTEEQL ENIADRTVQEADEDGDGAVSFVEFTKSLEKM DVEHKMSIRILK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-	l	USSN	location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		i .	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			İ	amino acid	of peptide	T=Threonine, V=Valine, W=Tryntophan
ł				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
	.]	1		peptide		/=possible nucleotide deletion, \=possible
060	10010	↓	-	sequence		nucleotide insertion
968	2318	A	8211	2	409	ISSCPHTAYEGSMSTLSNFTQTLEDVFRRIFIT
1	j]		ļ	1	YMDNWRQNTTAEQEALQAKVDAENFYYVIL
					1	YLMVMIGMFSFIIVAILVSTVKSKRREHSNDP
ļ	İ		1			YHQYIVEDWQEKYKSQILNLEESKATIHENIG AAGFKMSP
969	2319	A	8215	1	1938	GMPRSRGGRAAPGPPPPPPPPGQAPRWSRWR
İ	1	!		_		VPGRLLLLLPALCCLPGAARAAAAAAGAGN
	Į.	1				RAAVAVAVARADEAEAPFAGQNWLKSYGY
1	·]				LLPYDSRASALHSAKALQSAVSTMOOFYGIP
		İ				VTGVLDQTTIEWMKKPRCGVPDHPHLSRRRR
						NKRYALTGQKWRQKHITYSIHNYTPKVGELD
		ŀ	i			TRKAIRQAFDVWQKVTPLTFEEVPYHEIKSDR
		ł	l			KEADIMIFFASGFHGDSSPFDGEGGFLAHAYF
.	Ì					PGPGIGGDTHFDSDEPWTLGNANHDGNDLFL
						VAVHELGHALGLEHSSDPSAIMAPFYQYMET HNFKLPQDDLQGIQKIYGPPAEPLEPTRPLPTL
] .]			PVRRIHSPSERKHERQPRPPRPPLGDRPSTPGT
		1	f 1			KPNICDGNFNTVALFRGEMFVFKDRWFWRL
}			<u> </u>			RNNRVQEGYPMQIEQFWKGLPARIDAAYER
1	!	ł	i i	•		ADGRFVFFKGDKYWVFKEVTVEPGYPHSLG
						ELGSCLPREGIDTALRWEPVGKTYFFKGERY
						WRYSEERRATDPGYPKPITVWKGIPQAPQGA
1	}					FISKEGYYTYFYKGRDYWKFDNQKLSVEPGY
1			ĺ			PRNILRDWMGCNQKEVERRKERRLPQDDVDI
						MVTINDVPGSVNAVAVVIPCILSLCILVLVYTI FQFKNKTGPQPVTYYKRPVQEWV
970	2320	A	8216	1235	2223	SRLSLQFYVSFRRTGLFTCKLIVEIFFRNYMN
			•			DSLRTNVFVRFQPETIACACIYLAARALQIPLP
				•		TRPHWFLLFGTTEEEIQEICIETLRLYTRKKPN
					ļ	YELLEKEVEKRKVALQEAKLKAKGLNPDGTP
						ALSTLGGFSPASKPSSPREVKAEEKSPISINVK
		İ	i i			TVKKEPEDRQQASKSPYNGVRKDSKRSRNSR
1 1		' i	ł			SASRSRSRTRSRSRSHTPRRHYNNRRSRSGTY
}						SSRSRSRSRSHSESPRRHHNHGSPHLKAKHTR DDLKSSNRHGHKRKKSRSRSQSKSRDHSDAA
i i		}				KKHRHERGHHRDRRERSRSFERSHKSKHHGG
		j	į	j		SRSGHGRHRR
971	2321	Α	8217	3	3274	DCRLQAAMPTNFTVVPVEAHADGGGDETAE
)		ĺ				RTEAPGTPEGPEPERPSPGDGNPRENSPFLNN
]]	1		•	1	ľ	VEVEQESFFEGKNMALFEEEMDSNPMVSSLL.
1		j	i]	1	NKLANYTNLSQGVVEHEEDEESRRREAKAPR
1		ŀ		i	I	MGTFIGVYLPCLQNILGVILFLRLTWIVGVAG
	- 1				J	VLESFLIVAMCCTCTMLTAISMSAIATNGVVP AGGSYYMISRSLGPEFGGAVGLCFYLGTTFA
1	İ			İ		GAMYILGTIEIFLTYISPGAAIFQAEAAGGEAA
j l	1		- 1	1	!	AMLHNMRYYGTCTLVLMALVVFVGVKYVN
]	ĺ	[ſ		ĺ	KLALVFLACVVLSILAIYAGVIKSAFDPPDIPV
		i				CLLGNRTLSRRSFDACVKAYGIHNNSATSAL
	1		- 1		ĺ	WGLFCNGSQPSAACDEYFIQNNVTEIOGIPGA
1 1	ļ	ł		ļ	ļ	ASGVFLENLWSTYAHAGAFVEKKGVPSVPV
						AEESRASTLPYVLTDIAASFTLLVGIYFPSVTG
]]	- 1	-			. [IMAGSNRSGDLKDAQKSIPTGTILAIVTTSFIY
	1	1			ľ	LSCIVLFGACIEGVVLRDKFGEALQGNLVIGM LAWPSPWVIVIGSFFSTCGAGLQTLTGAPRLL
		ł	'			QAIARDGIVPFLQVFGHGKANGEPTWALLLT
	- 1	- 1				VLICETGILIASLDSVAPILSMFFLMCYLFVNL
}	I	-		ļ	j	ACAVQTLLRTPNWRPRFKFYHWTLSFLGMSL
		l	- 1			CLALMFICSWYYALSAMLIAGCIYKYIEYRG
		1			. 1	AEKEWGDGIRGLSLNAARYALLRVEHGPPHT
						KNWRPQVLVMLNLDAEQAMKHPRLLSFTSQ

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LKAGKGLTIVGSVLEGTYLDKHMEAQRAEE NIRSLMSTEKTKGFCQLVVSSSLRDGMSHLIQ SAGLGGLKHNTVLMAWPASWKQEDNPFSW KNFVDTVRDTTAAHQALLVAKNVDSFPQNQ ERFGGGHIDVWWIVHDGGMLMLLPFLLRQH KVWRKCRMRIFTVAQVDDNSIQMKKDLQMF LYHLRISAEVEVVEMVENDISAFTYERTLMM EQRSQMLKQMQLSKNEQEREAQLIHDRNTAS HTAAAARTQAPPTPDKVQMTWTREKLIAEK YRSRDTSLSGFKDLFSMKPDQSNVRRMHTAV KLNGVVLNKSQDAQLVLLNMPGPPKNRQGD ENYMEFLEVLTEGLNRVLLVRGGGREVITIYS
972	2322	A	8224	701	246	TSRRVTMKFNPFVTSDRSKNRKRHFNAPSHV RRKIMSSPLSKELRQKYNVRSMPIRKDDEVQ VVRGHYKGQQIGKVVQVYRKKYVIYIERVQ REKANGTTVHVGIHPSKVVITRLKLDKDRKKI LERKAKSRQVGKEKGKYKEELIEKMQE
973	2323	A	8237	279	4610	GCPHAGGKGRVPTGGLTGGRTWSPSAAPRSC PRPGPTPAPGAMDKLPPSMRKRLYSLPQQVG AKAWIMDEEDAEEEGAGGRQDPSRRSIRLR PLPSSPSAAAAGGTESRSSALGAADSEGPARG AGKSSTNGDCRRFRGSLASLGSRGGSGGTG SGSSHGHLHDSAEERRLIAEGDASPGEDRTPP GLAAEPERPGASAQPAASPPPPQQPPQPASAS CEQPSVDTAIKVEGGAAAGDQILPEAEVRLG QAGFMQRQFGAMLQPGVNKFSLRMFGSQKA VEREQERVKSAGFWIIHPYSDFRFYWDLTML LLMVGNLIIIPVGITFFKDENTTPWIVFNVVSD TFFLIDLVLNFRTGIVVEDNTEIILDPQRIKMK YLKSWFMVDFISSIPVDYIFLIVETRIDSEVYK TARALRIVRFTKILSLLRLLRLSRLIRYIHQWE EIFHMTYDLASAVVRIVNLIGMMLLLCHWDG CLQFLVPMLQDFPDDCWVSINNMVNNSWGK QYSYALFKAMSHMLCIGYGRQAPVGMSDV WLTMLSMIVGATCYAMFIGHATALIQSLDSS RRQYQEKYKQVEQYMSFHKLPPDTRQRIHD YYEHRYQGKMFDEESILGELSEPLREEIINFNC RKLVASMPLFANADPNFVTSMLTKLRFEVFQ PGDYIIREGTIGKKMYFIQHGVVSVLTKGNKE TKLADGSYFGEICLLTRGRRTASVRADTYCR LYSLSVDNFNEVLEEYPMMRRAFETVALDRL DRIGKKNSILLHKVQHDLNSGVFNYQENEIIQ QIVQHDREMAHCAHRVQAAASATPTPTPVIW TPLIQAPLQAAAATTSVAIALTHHPRLPAAIFR PPPGSGLGNLGAGGYPRHLKRLQSLIPSALGS ASPASSPSQVDTPSSSSFHIQQLAGFSAPAGLS PLLPSSSSSPPPGACGSPSAPTPSAGVAATTIA GFGHFHKALGGSLSSSDSPLLTPLQPGARSPQ AAQPSPAPPGARGGLGLPEHFLPPPPSSRSPS SPGQLGQPPGELSLGLATGPLSTPETPPRQPEP PSLVAGASGGASPVGFTPRGGLSPPGHSPGPP RTFPSAPPRASGSHGSLLLPPASSPPPPQVPQR RGTPPLTPGRLTQDLKLISASQPALPQDGAQT LRRASPHSSGESMAAFPLFPRAGGGSGGSGS GGLGPPGRPYGAIPGOHVTLPRKTSSGSLPPP LSLFGARATSSGGPPLTAGPQREPGARPEPVR SKLPSNL
974	2324	A	8247	279	468	EYKQWERRFLSCQNRNDLGYGKPRKGGGLL LVPVKDASRICSLTYLLGSHWNNLVVRSPVL G

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
975	2325	A	8249	62	1571	LVALKNWKPKGTNIPAPQSPVFGEAVSGVYM MTKVLGMAPVLGPRPPQEQVGPLMVKVEEK EEKGKYLPSLEMFRQRFRQFGYHDTPGPREA LSQLRVLCCEWLRPEIHTKEQILELLVLEQFLT ILPQELQAWVQEHCPESAEEAVTLLEDLEREL DEPGHQVSTPPNEQKPVWEKISSSGTAKESPS SMQPQPLETSHKYESWGPLYIQESGEEQEFAQ DPRKVRDCRLSTQHEESADEQKGSEAEGLKG DIISVIIANKPEASLERQCVNLENEKGTKPPLQ EAGSKKGRESVPTKPTPGERRYICAECGKAFS NSSNLTKHRRTHTGEKPYVCTKCGKAFSHSS NLTLHYRTHLVDRPYDCKCGKAFGQSSDLLK HQRMHTEEAPYQCKDCGKAFSGKGSLIRHYR IHTGEKPYQCNECGKSFSQHAGLSSHQRLHT GEKPYKCKECGKAFNHSSNFNKHHRIHTGEK PYWCHHCGKTFCSKSNLSKHQRVHTGEGEA P
976	2326	A	8257	298	7086	GNMACWPOLRLLLWKNLTFRRQTCQLLLE VAWPLFIFLILISVRLSYPPYEQHECHFPNKAM PSAGTLPWVQGIICNANNPCFRYPTPGEAPGV VGNFNKSIVARLFSDARRLLLYSQKDTSMKD MRKVLRTLQQIKKSSSNLKLQDFLVDNETFS GFLYHNLSLPKSTVDKMLRADVILHKVFLQG YQLHLTSLCNGSKSEEMIQLGDQEVSELCGLP REKLAAAERVLRSNMDILKPILRTLNSTSPFPS KELAEATKTLLHSLGTLAQELFSMRSWSDMR QEVMFLTNVNSSSSSTQIYQAVSRIVCGHPEG GGLKIKSLNWYEDNNYKALFGGNGTEEDAE TFYDNSTTPYCNDLMKNLESSPLSRIIWKALK PLLVGKILYTPDTPATRQVMAEVNKTFQELA VFHDLEGMWEELSPKIWTFMENSQEMDLVR MLLDSRDNDHFWEQQLDGLDWTAQDIVAFL AKHPEDVQSSNGSVYTWREAFNETNQAIRTIS RFMECVNLNKLEPIATEVWLINKSMELLDER KFWAGIVFTGITPGSIELPHHVKYKIRMGIDN VERTNKIKDGYWDPGPRADPFEDMRYVWGG FAYLQDVVEQAIIRVLTGTEKKTGVYMQQMP YPCYVDDIFLRVMSRSMPLFMTLAWIYSVAV IIKGIVYEKEARLKETMRIMGLDNSILWFSWFI SSLIPLLVSAGLLVVILKLGNLLPYSDPSVVFV FLSVFAVVTILQCFLISTLFSRANLAAACGGII YFTLYLPYVLCVAWQDYVGFTLKIFASLLSP VAFGFGCEYFALFEEQGIGVQWDNLFESPVE EDGFNLTTSVSMMLFDTFLYGVMTWYIEAVF PGQYGIPRPWYFPCTKSYWFGEESDEKSHPGS NQKRISEICMEEEPTHLKLGVSIQNLVKVYRD GMKVAVDGLALNFYEGQITSFLGHNGAGKT TTMSILTGLFPPTSGTAYII.GKDTRSEMSTIRQ NLGVCPQHNVLFDMLTVEEHIWFYARLKGLS EKHVKAEMEQMALDVGLPSSKLKSKTSQLS GGMQRKLSVALAFVGGSKVVILDEPTAGVDP YSRRGIWELLLKYRQGRTILSTHHMDEADVL GDRIAIISHGKLCCVGSSLFLKNQLGTGYYLT LVKKDVESSLSSCRNSSSTVSYLKKEDSVSQS SSDAGLGSDHESDTLTIDVASISNLIRKHVYSEA RLVEDIGHELTYVLPYEAAKEGAFVELFHEID DRLSDLGISSYGISETTLEEIFLKVAEESGVDA ETSDGTLPARRNRAFGDKQSCLRPFTEDDA ADPNDSDIDPESRETDLLSGMDGKGSYQVKG WKLTQQQFVALLWKRLLIARRSRKGFFAQIV

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
077	3207		P2.60			LPAVFVCIALVFSLIVPPFGKYPSLELQPWMY NEQYTFVSNDAPEDTGTLELLNALTKDPGFG TRCMEGNPIPDTPCQAGEEEWTTAPVPQTIM DLFQNGNWTMQNPSPACQCSSDKIKKMLPV CPPGAGGLPPPQRKQNTADILQDLTGRNISDY LVKTYVQIIAKSLKNKIWVNEFRYGGFSLGVS NTQALPPSQEVNDATKQMKKHLKLAKDSSA DRFLNSLGRFMTGLDTRNNVKVWFNNKGW HAISSFLNVINNAILRANLQKGENPSHYGITAF NHPLNLTKQQLSEVAPMTTSVDVLVSICVIFA MSFVPASFVVFLIQERVSKAKHLQFISGVKPVI YWLSNFVWDMCNYVVPATLVIIFICFQQKSY VSSTNLPVLALLLLLYGWSITPLMYPASFVFK IPSTAYVVLTSVNLFIGINGSVATFVLELFTDN KLNNINDILKSVFLIFPHFCLGRGLIDMVKNQ AMADALERFGENRFVSPLSWDLVGRNLFAM AVEGVVFFLITVLIQYRFFIRPRPVNAKLSPLN DEDEDVRRERQRILDGGGQNDILEIKELTKIY RRKRKPAVDRICVGIPPGECFGLLGVNGAGK SSTFKMLTGDTTVTRGDAFLNRNSILSNIHEV HQNMGYCPQFDAITELLTGREHVEFFALLRG VPEKEVGKVGEWAIRKLGLVKYGEKYAGNY SGGNKRKLSTAMALIGGPPVVFLDEPTTGMD PKARRFLWNCALSVVKEGRSVVLTSHSMEEC EALCTRMAIMVNGRFRCLGSVQHLKNRFGD GYTIVVRIAGSNPDLKPVQDFFGLAFPGSVPK EKHRNMLQYQLPSSLSSLARIFSILSQSKYRLH IEDYSVSQTTLDQVFVNFAKDQSDDDHLKDL SLHKNOTVVDVAVLTSFLQDEKVKESYV
977	2327	A	8260	3	1567	IPGSTISFSLCFIFPPCVPTMVRKPVVSTISKGG YLQGNVNGRLPSLGNKEPPGQEKVQLKRKV TLLRGVSIIIGTIIGAGIFISPKGVLQNTGSVGM SLTIWTVCGVLSLFGALSYAELGTTIKKSGGH YTYILEVFGPLPAFVRVWVELLIIRPAATAVIS LAFGRYILEPFFIQCEIPELAIKLITAVGITVVM VLNSMSVSWSARIQIFLTFCKLTAILIIIVPGV MQLIKGQTQNFKDAFSGRDSSITRLPLAFYYG MYAYAGWFYLNFVTEEVENPEKTIPLAICISM AIVTIGYVLTNVAYFTTINAEELLLSNAVAVT PSERLGNFSLAVPIFVALSCFGSMNGGVFAV SRLFYVASREGHLPEILSMIHVRKHTPLPAVIV LHPLTMIMLFSGDLDSLLNFLSFARWLFIGLA VAGLIYLRYKCPDMHRPFKVPLFIPALFSFTC LFMVALSLYSDPFSTGIGFVITLTGVPAYYLFII WDKKPRWFRIMSEKITRTLQIILEVVPEEDKL
978	2328	A	8261	2	2165	RGGSLRCVLGKLLGQLLCFQSERCVRFPEGLL RHRGCGLLSSRLSAGKPPLRTSFFGSWGVLPP LADAASMSGVRAVRISIESACEKQVHEVGLD GTETYLPPLSMSQNLARLAQRIDFSQGSGSEE EEAAGTEGDAQEWPGAGSSADQDDEEGVVK FQPSLWPWDSVRNNLRSALTEMCVLYDVLSI VRDKKFMTLDPVSQDALPPKQNPQTLQLISK KKSLAGAAQILLKGAERLTKSVTENQENKLQ RDFNSELLRLRQHWKLRKVGDKILGDLSYRS AGSLFPHHGTFEVIKNTDLDLDKKIPEDYCPL DVQIPSDLEGSAYIKVSIQKQAPDIGDLGTVN LFKRPLPKSKPGSPHWQTKLEAAQNVLLCKEI FAQLSREAVQIKSQVPHIVVKNQIISQPFPSLQ LSISLCHSSNDKKSQKFATEKQCPEDHLYVLE HNLHLLIREFHKQTLSSIMMPHPASAPFGHKR

SEQ I	D SEO ID	Met	SEQ	Predicted	Dradiated and	Amino said regues of AmAlanias C. O.
NO: 0		hod	ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	,,,,,,	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide		1	USSN	location		I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496		to last amino	
uence	1 .		914	ng to first	acid residue	M=Methionine, N=Asparagine, P=Proline,
ucucc	- }	J	714	amino acid		Q=Glutamine, R=Arginine, S=Serine,
1	1	1		residue of	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1		1			sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1		ł	ŀ	peptide	1	/=possible nucleotide deletion, \=possible
			 	sequence		nucleotide insertion
	i	ļ		1		MRLSGPQAFDKNEINSLQSSEGLLEKIIKQAK
1	i	ł		1	1	HIFLRSRAAATIDSLASRIEDPQIQAHWSNIND
	1	Ì			1	VYESSVKVLITSQGYEQICKSIQLQLNIGVEQI
ļ	- {					RVVHRDGRVITLSYQEQELQDFLLSQMSQHQ
		1				VHAVQQLAKVMGWQVLSFSNHVGLGPIESIG
1	- 1	ł		•		NASAITVASPSGDYAISVRNGPESGSKIMVQF
	l					PRNQCKDLPKSDVLQDNKWSHLRGPFKEVQ
						WNKMEGRNFVYKMELLMSALSPCLL
979	2329	Α	8289	2	1053	FVWNPRGGRKRRRQAAVTQAATRASGTPSP
	- 1	į				RDGTMTQGKLSVANKAPGTEGQQQVHGEKK
		l		ļ		EAPAVPSAPPSYEEATSGEGMKAGAFPPAPTA
1	ł	ľ	1			VPLHPSWAYVDPSSSSSYDNGFPTGDHELFTT
1	1	İ	1			FSWDDQKVRRVFVRKVYTILLIQLLVTLAVV
						ALFTFCDPVKDYVQANPGWYWASYAVFFAT
		ł				YLTLACCSGPRRHFPWNLILLTVFTLSMAYLT
1	ľ	l	1			GMLSSYYNTTSVLLCLGITALVCLSVTVFSFQ
1		1				TKFDFTSCQGVLFVLLMTLFFSGLILAILLPFO
						YVPWLHAVYAALGAGVFTLFLALDTQLLMG
ſ		i	1			NRRHSLSPEEYIFGALNIYLDIIYIFTFFLOLFG
}	ŀ	1				TNRE
980	2330	Α	8305	59	857	ASQLPDYSISPPSLPPRISFHPSPTLARVAMAEP
						SEATQSHSISSSSFGAEPSAPGGGGSPGACPAL
1		ľ	1			GTKSCSSSCAVHDLIFWRDVKKTGFVFGTTLI
1						MLLSLAAFSVISVVSYLILALLSVTISFRIYKSV
1		ł				IQAVQKSEEGHPFKAYLDVDITLSSEAFHNY
	1	ľ	1			
	l l			,		MNAAMVHINRALKLIIRLFLVEDLVDSLKLA
	·					VFMWLMTYVGAVFNGITLLILAELLIFSVPIV
Ĭ			1 1	. i		YEKYKTQIDHYVGIARDQTKSIVEKIQAKLPG
981	2331	A	8308	186	1227	IAKKKAE
701	2331	A	8308	100	1337	TRMSRHEGVSCDACLKGNFRGRRYKCLICYD
	1					YDLCASCYESGATTTRHTTDHPMQCILTRVD
1			1 1	ł		FDLYYGGEAFSVEQPQSFTCPYCGKMGYTET
			1		1	SLQEHVTSEHAETSTEVICPICAALPGGDPNH
1						VTDDFAAHLTLEHRAPRDLDESSGVRHVRR
ĺ						MFHPGRGLGGPRARRSNMHFTSSSTGGLSSS
	j			ļ		QSSYSPSNREAMDPIAELLSQLSGVRRSAGGQ
1				1		LNSSGPSASQLQQLQMQLQLERQHAQAARQ
1				ĺ		QLETARNATRRTNTSSVTTTITQSTATTNIAN
i				i		TESSQQTLQNSQFLLTRLNDPKMSETERQSM
1	1					ESERADRSLFVQELLLSTLVREESSSSDEDDR
1			1			GEMADFGAMGCVDIMPLDVALENLNLKESN
						KGNEPPPPPL
982	2332	A	8315	1	1004	GSTHASADAWAQWFCTEALVMGAPVWYLV
1			ļ ļ		<u>. </u>	AAALLVGFILFLTRSRGRAASAGQEPLHNEEL
1	1 1			i	i	AGAGRVAQPGPLEPEEPRAGGRPRRRRDLGS
				Į.	1	RLQAQRRAQRVAWAEADENEEEAVILAQEE
1	j j		, 1	j	J	EGVEKPAETHLSGKIGAKKLRKLEEKOARKA
1				ĺ	ļ	QREAEEAEREERKRLESQREAEWKKEEERLR
1		ļ	.	1	İ	LEEEQKEEEERKAREEQAQREHEEYLKLKEA
1				i	1	FVVEEEGVGETMTEEQSQSFLTEFINYIKOSK
1	1 1				1	VVLLEDLASQVGLRTQDTINRIQDLLAEGTIT
	[]			1	1	GVIDDRGKFIYITPEELAAVANFIRORGRVSIA
	1			1	l	ELAQASNSLIAWGRESPAQAPA
983	2333	Α	8320	244	1420	RRRWRARGGLVPTLAWAEATGAYVPGRDKP
			3323		.720	
					1	DLPTWKRNFRSALNRKEGLRLAEDRSKDPHD
l					1	PHKIYEFVNSGVGDFSQPDTSPDTNGGGSTSD
]]	1	ļ	J	ī	
1					1	TQEDILDELLGNMVLAPLPDPGPPSLAVAPEP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, B=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion EVGDRTLPGWPVTLPDPGMSLTDRGVMSYV RHVLSCLGGGLALWRAGQWLWAQRLGHCH TYWAVSEELLPNSGHGPDGEVPKDKEGGVF DLGPPIVGSLGPPDLITFTEGSGRSPRYALWFC VGESWPQDQPWTKRLVMVKVVPTCLRALVE MARVGGASSLENTVDLHISNSHPLSLTSDQY KAYLQDLVEGMDFQGPGES
984	2334	A	8321	1	1243	ANMAPVEHVVADAGAFLRHAALQDIGKNIY TIREVVTEIRDKATTRRLLAVLPYELRFKEPLPE YVRLVTEFSKKTGDYPSI SATDIQVLALTYQL EAEFVGVSHLKQEPQKVKVSSSIQHPETPLHIS GFHLPYKPKPPQETEKGHSACEPENLEFSSFM FWRNPLPNIDHELQELLIDRGEDVPSEEEEEE NGFEDRKDDSDDDGGGWITPSNIKQIQQELE QCDVPEDVRVGCLTTDFAMQNVLLQMGLHV LAVNGMLIREARSYILRCHGCFKTTSDMSRV FCSHCGNKTLKKVSVTVSDDGTLHMHFSRNP KVLNPRGLRYSLPTPKGGKYAINPHLTEDQRF PQLRLSQKARQKTNVFAPDYIAGVSPFVENDI SSRSATLQVRDSTLGAGRRRLNPNASRKKFV KKR
985	2335	Α	8322	352	529	RRNNIRQFIMKVCISGQARWLTPVVPVLWET EAGRSLELKSLRPAWATWGNPISTKINK
986	2336	A	8325	89	1172	KMNPTDIADTTLDESIYSNYYLYESIPKPCTKE GIKAFGELFLPPLYSLVFVFGLLGNSVVVLVL FKYKRLRSMTDVYLLNLAISDLLFVFSLPFWG YYAADQWVFGLGLCKMISWMYLVGFYSGIF FVMLMSIDRYLAIVHAVFSLRARTLTYGVITS LATWSVAVFASLPGFLFSTCYTERNHTYCKT KYSLNSTTWKVLSSLEINILGLVIPLGIMLFCY SMIIRTLQHCKNEKKNKAVKMIFAVVVLFLG FWTPYNIVLFLETLVELEVLQDCTFERYLDYA IQATETLAFVHCCLNPIIYFFLGEKFRKYILQL FKTCRGLFVLCQYCGLLQIYSADIPSSSYTQS TMDHDLHDAL
987	2337	A	8326	3	470	SLSAMRFLAATFLLLALSTAAQAEPVQFKDC GSVDGVIKEVNVSPCPTQPCQLSKGQSYSVN VTFTSNIQSKSSKAVVHGILMGVPVPFPIPEPD GCKSGINCPIQKDKTYSYLNKLPVKSEYPSIK LVVEWQLQDDKNQSLFCWEIPVQIVSHL
988	2338	A	8335	1205	323	VIKMALAARLLPOFLHSRSLPCGAVRLRTPA VAEVRLPSATLCYFCRCRLGLGAALFPRSAR ALAASALPAQGSRWPVLSSPGLPAAFASFPAC PQRSYSTEEKPQQHQKTKMIVLGFSNPINWV RITIKAFLIWAYFDKEFSITEFSEGAKQAFAH VSKLLSQCKFDLLEELVAKEVLHALKEKVTS LPDNHKNALAANIDEIVFTSTGDISIYYDEKG RKFVNILMCFWYLTSANIPSETLRGASVFQVK LGNQNVETKQLLSASYEFQREFTQGVKPDWT IARIEHSKLLE
989	2339	Α	8349	67	185	MSGFIHQLLIQNLFCVYHTRLKTSQGLCLLSL
990	2340	A	8361	210	1115	KSLHPMS ASPFLRPQGHDSGEREPFSQTPGLMQPFSIPVQ ITLQGSRRRQGRTAFPASGKKRETDYSDGDPL DVHKRLPSSTGEDRAVMLGFAMMGFSVLMF FLLGTTILKPFMLSIQREESTCTAIHTDIMDDW LDCAFTCGVHCHGQGKYPCLQVFVNLSHPG QKALLHYNEEAVQINPKCFYTPKCHQDRNDL LNSALDIKEFFDHKNGTPFSCFYSPASQSEDVI

CEO ID	CEOID	Mat	LEEO	Dunding	N	E A
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	100100		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
Lacito	ĺ	[714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
	ŀ			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	Sequence	/=possible nucleotide deletion, \=possible
	İ			sequence	•	nucleotide insertion
		— —	 	- coquesion		LIKKYDQMAIFHCLFWPSLTLLGGALIVGMV
					İ	RLTQHLSLLCEKYSTVVRDEVGGKVPYIEQH
			ł		1	QFKLCIMRRSKGRAEKS
991	2341	A	8369	9	921	SSVVEFSALSVSMACLSPSQLQKFQQDGFLVL
Ì						EGFLSAEECVAMQQRIGEIVAEMDVPLHCRT
		ļ	1			EFSTQEEEQLRAQGSTDYFLSSGDKIRFFFEK
 		1	ł			GVFDEKGNFLVPPEKSINKIGHALHAHDPVFK
						SITHSFKVQTLARSLGLQMPVVVQSMYIFKQP
į						HFGGEVSPHQDASFLYTEPLGRVLGVWIAVE
						DATLENGCLWFIPGSHTSGVSRRMVRAPVGS
l .		i				APGTSFLGSEPARDNSLFVPTPVQRGALVLIH
		l l			i	GEVVHKSKQNLSDRSRQAYTFHLMEASGTT
]				WSPENWLQPTAELPFPQLYT
992	2342	A	8370	906	4	MALSGNCSRYYPREQGSAVPNSFPEVVELNV
						GGQVYFTRHSTLISIPHSLLWKMFSPKRDTAN
			•	1		DLAKDSKGRFFIDRDGFLFRYILDYLRDRQVV
		ļ				LPDHFPEKGRLKREAEYFQLPDLVKLLTPDEI
						KQSPDEFCHSDFEDASQGSDTRICPPSSLLPAD
						RKWGFITVGYRGSCTLGREGQADAKFRRVPR
		l	1			ILVCGRISLAKEVFGETLNESRDPDRAPERYTS
						RFYLKFKHLMGAPASNFILGFWGLGQNQDK
						HPVNIYLQQRSVIRPDLTSKKAGDLKGKGDA
						QEVSRRRRWLGDPEHL
993	2343	A	8379	1	2794	MRMQRHKNDTMDFGDSGKRIGGGVLCLLHQ
						SNTSFIKLNNNGFEDIVIVIDPSVPEDEKIIEQIE
						DMVTTASTYLFEATEKRFFFKNVSILIPENWK
						ENPQYKRPKHENHKHADVIVAPPTLPGRDEP
	•			,		YTKQFTECGEKGEYIHFTPDLLLGKKQNEYG
1						PPGKLFVHEWAHLRWGVFDEYNEDQPFYRA
]			ŀ			KSKKIEATRCSAGISGRNRVYKCQGGSCLSRA
			1			CRIDSTTKLYGKDCQFFPDKVQTEKASIMFM
1						QSIDSVVEFCNEKTHNQEAPSLQNIKCNFRST
i 1			' I			WEVISNSEDFKNTIPMVTPPPPPVFSLLKIRQRI
1						VCLVLDKSGSMGGKDRLNRMNQAAKHFLLQ TVENGSWVGMVHFDSTATIVNKLIQIKSSDER
!						NTLMAGLPTYPLGGTSICSGIKYAFQVIGELH
1						SQLDGSEVLLLTDGEDNTASSCIDEVKQSGAI
						VHFIALGRAADEAVIEMSKITGGSHFYVSDEA
Į						QNNGLIDAFGALTSGNTDLSQKSLQLESKGLT
						LNSNAWMNDTVIIDSTVGKDTFFLITWNSLPP
j						SISLWDPSGTIMENFTVDATSKMAYLSIPGTA
{					l	KVGTWAYNLQAKANPETLTTTVTSRAANSSV
					ĺ	PPITVNAKMNKDVNSFPSPMIVYAEILOGYVP
						VLGANVTAFIESONGHTEVLELLDNGAGADS
			'	·	1	FKNDGVYSRYFTAYTENGRYSLKVRAHGGA
						NTARLKLRPPLNRAAYIPGWVVNGEIEANPP
[[- [1	j	RPEIDEDTQTTLEDFSRTASGGAFVVSQVPSL
		1			ļ	PLPDQYPPSQITDLDATVHEDKIILTWTAPGD
[NFDVGKVQRYIIRISASILDLRDSFDDALQVN
						TTDLSPKEANSKESFAFKPENISEENATHIFIAI
[1		i		KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPT
						PTPTPTPTPDKSHNSGVNISTLVLSVIGSVVIV
L						NFILSTTI
994	2344	A	8385	231	644	INSSPRTGRDHQELNLHTERDSRSQRAVLKIP
<u> </u>		1				RQNPGIFYWIFLPSRSHSASHGSRQRQVSCQG
			}	1		TQDEILKMRNTFAELKNSLEALSSRMDQAEE
}						RIGTQAGVQWRDHGSLQPQPPEFKQCFHLSL
						PSSWDYRACLS
995	2345	Ā	8390	194	3421	AWRKSSVVPPRGTRRGEKSDQDKSGQKNKR

SEQ ID SEQ ID Not of nucleotide eotide sequence Predicted end nucleotide ID NO: of nucleotide Sequence Sequ
nucleotide sequence uenc
sequence Sequence
uence 914 ng to first amino acid residue of peptide sequence 914 ng to first amino acid residue of peptide sequence 925 94 95 96 97 97 97 97 97 97 97 97 97
amino acid residue of peptide sequence
residue of peptide sequence Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \text{\text{\text{Possible}}} possible nucleotide deletion, \text{\text{\text{\text{Possible}}}} possible nucleotide deletion, \text{\text{\text{\text{\text{Possible}}}} possible nucleotide deletion, \text{\t
A-possible nucleotide deletion, \-possible nucleotide insertion
sequence DFLSMKQSPALAPEERCRRAGSPKPVLRADD NNMGNGCSQKLATANLLRFLLLVLIPCICALN LLLEILLSYVGTLQKVYFKSNGSEPLVTDGEI QGSDVILTNTTYNQSTVVSTAHPDQHVPAWT TDASLPGDQSHRNTSACMNITHSQCQMLPYF ATLTPLLSVVRNMEMEKFLKFFTYLHRLSCY QHIMLFGCTLAFPECIIDGDDSHGLPCRSFCE AAKEGCESVLGMVNYSWPDFLRCSQFRNQT ESSNVSRICFSPQENGKQLLCGRGENFLCAS GICIPGKLQCNGYNDCDDWSDEAHCNCSENI FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCLAMEWVCDGDHDCVI KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLFYNSTSYPNYTGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSG CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQG CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHIGGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
DFLSMKQSPALAPEERCRAGSPKPVLRADD NNMGNGCSQKLATANLLRFLLLVLIPCICALY LLLEILLSYVGTLQKVYFKSNGSEPLVTDGEI QGSDVLTNTTYNQSTVVSTAHPDQHVPAWT TDASLPGDQSHRNTSACMNITHSQCQMLPYF ATLTPLLSVVRMEMEKFLKFFTYLHRLSCY QHIMLFGCTLAFPECIIDGDSHGLLPCRSFCE AAKEGCESVLGMVNYSWPDFLRCSQFRNQT ESSNVSRICFSPQQENGKQLLCGRGENFLCAS GICIPGKLQCNGYNDCDDWSDEAHCNCSENI FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVI KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSBEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSGG CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQG QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRNLGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
NNMGNGCSQKLATANLLRFLLLVLIPCICALY LLLEILLSYVGTI-QKVYFKSNGSEPLVTDGEI QGSDVIL.TNTTYNQSTVVSTAHPDQHVPAWT TDASLPGDQSHRNTSACMNITHSQCQMLPYF ATLTPLLSVVRNMEMEKFLKFFTYLHRLSCY QHIMLFGCTLAFPECIIDGDDSHGLLPCRSFCE AAKEGCESVLGMVNYSWPDFLRCSQFRNQT ESSNVSRICFSPQQENGKQLLCGRGENFLCAS GICPGGKLQCNGYNDCDDWSDEAHCNCSENI FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVI KSDEVNCSCHSQGLVECRNQQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSG CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRLLGGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
LLLEILLSYVGTLQKVYFKSNGSEPLVTDGEI QGSDVILTNTIYNQSTVVSTAHPDQHVPAWT TDASLPGDQSHRNTSACMNITHSQCQMLPYF ATLTPLLSVVRNMEMEKFLKFTIYLHRLSCY QHIMLFGCTLAFPECIIDGDDSHGLLPCRSFCE AAKEGCESVLGMVNYSWPDFLRCSQFRNQT ESSNVSRICFSPQQENGKQLLCGRGENFLCAS GICIPGKLQCNGYNDCDDWSDEAHCNCSENI FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVI KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSG CVLASRRCDGQADCDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
QGSDVILTNTTYNQSTVVSTAHPDQHVPAWT TDASLPGDQSHRNTSACMNITHSQCQMLPYF ATLTPLLSVVRNMEMEKFLKFITYLHRLSCY QHIMLFGCTLAFPECIIDGDDSHGLLPCRSFCE AAKEGCESVLGMVNYSWPDFLRCSQFRNQT ESSNVSRICFSPQQENGKQLLCGRGENFLCAS GICIPGKLQCNGYNDCDDWSDEAHCNCSENI FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVI KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPP CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSGG CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
TDASLPGDQSHRNTSACMNITHSQCQMLPYF ATLTPLLSVVRNMEMEKFLKFFTYLHRLSCY QHIMLFGCTLAFPECIIDGDDSHGLLPCRSFCE AAKEGCESVLGMVNYSWPDFLRCSQFRNQT ESSNVSRICFSPQQENGKQLLCGRGENFLCAS GICPGKLQCNGYNDCDDWSDEAHCNCSENI FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVI KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSGG CVLASRRCDGQADCDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
ATLTPLLSVVRNMEMEKFLKFFTYLHRLSCY QHIMLFGCTLAFPECIIDGDDSHGLLPCRSFCE AAKEGESVLGMVNYSWPDFLRCSQFRNQT ESSNVSRICFSPQQENGKQLLCGRGENFLCAS GICPGKLQCNGYNDCDDWSDEAHCNCSENI FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVI KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSG CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
QHIMLFGCTLAFPECIIDGDDSHGLLPCRSFCE AAKEGCESVLGMVNYSWPDFLRCSQFRNQT ESSNVSRICFSPQQENGKQLLCGRGENFLCAS GICIPGKLQCNGYNDCDDWSDEAHCNCSENI FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVI KSDEVNCSCHSQGLVECRNQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSGC CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
AAKEGCESVLGMVNYSWPDFLRCSQFRNQT ESSNVSRICFSPQQENGKQLLCGRGENFLCAS GICIPGKLQCNGYNDCDDWSDEAHCNCSENI FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVI KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSQ CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQELLSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQG CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
ESSNVSRICFSPQQENGKQLLCGRGENFLCAS GICIPGKLQCNGYNDCDDWSDEAHCNCSENI FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVU KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSGG CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQELSQLACKQMGLGEPSVTKLIQE QEKEPRWLTHSNWESLNGTTLHELLVNGGG CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
GICIPGKLQCNGYNDCDDWSDEAHCNCSENI FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVI KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSQC CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQELSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESJLNGTTLHELLVNGQG CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
FHCHTGKCLNYSLVCDGYDDCGDLSDEQNC DCNPTTEHRCGDGRCIAMEWVCDGDHDCVI KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSBEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSGG CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
DCNPTTEHRCGDGRCIAMEWVCDGDHDCVI KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSGG CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
KSDEVNCSCHSQGLVECRNGQCIPSTFQCDG DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSG CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGGS CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
DEDCKDGSDEENCSVIQTSCQEGDQRCLYNP CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSGC CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESSLNGTTLHELLVNGGS CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
CLDSCGGSSLCDPNNSLNNCSQCEPITLELCM NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSG FPEENSDNQTCLMPDEYVEECSPSHFKCRSGG CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQELLSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGGG CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
NLPYNSTSYPNYFGHRTQKEASISWESSLFPA LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSGC CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQG CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
LVQTNCYKYLMFFSCTILVPKCDVNTGEHIPF CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSGG CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
CRALCEHSKERCESVLGIVGLQWPEDTDCSQ FPEENSDNQTCLMPDEYVEECSPSHFKCRSGG CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGG TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
FPEENSDNQTCLMPDEYVEECSPSHFKCRSGG CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
CVLASRRCDGQADCDDDSDEENCGCKERDL WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGGE CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
WECPSNKQCLKHTVICDGFPDCPDYMDEKN CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGGY CESRSKISLLCTKQDCGRRPAARMNKRILGGY TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
CSFCQDDELECANHACVSRDLWCDGEADCS DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTILHSNWESLINGTTLHELLVNGQG; CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
DSSDEWDCVTLSINVNSSSFLMVHRAATEHH VCADGWQELLSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQG CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
VCADGWQEILSQLACKQMGLGEPSVTKLIQE QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
QEKEPRWLTLHSNWESLNGTTLHELLVNGQS CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
CESRSKISLLCTKQDCGRRPAARMNKRILGGI TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
TSRPGRWPWQCSLQSEPSGHICGCVLIAKKW VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
VLTVAHCFEGRENAAVWKVVLGINNLDHPS VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
VFMQTRFVKTIILHPRYSRAVVDYDISIVELSE
] DISETGYVRPVCLPNPEQWLEPDTYCYITGW
GHMGNKMPFKLQEGEVRIISLEHCQSYFDMK
TITTRMICAGYESGTVDSCMGDSGGPLVCEK
PGGRWTLFGLTSWGSVCFSKVLGPGVYSNVS
YFVEWIKRQIYIQTFLLN
996 2346 A 8392 199 3085 KVILSSEMSKTNKSKSGSRSSRSRSRSRSRS
FSKSRSRSRSRSRSRSRSRSRSRSRSSSSPAHN
RERNHPRVYQNRDFRGHNRGYRRPYYFRGR
NRGFYPWGQYNRGGYGNYRSNWQNYRQAY
SPRRGRSRSPKRRSPSPRSRSHSRNSDKSSS
DRSRRSSSRSSNHSRVESSKRKSAKEKKSSS
KDSRPSQAAGDNQGDEVKEQIFSGGTSQDTK
ASESSKPWPDATYGTGSASRASAVSELSPRER
SPALKSPLQSVVVRRRSPRPSPVKPSPPLSST
SQMGSTLPSGAGYQSGTHQGQFDHGSGSLSP
SKKSPVGKSPPSTGSTYGSSQKEESAASGGAA
YTKRYLEEQKTENGKDKEQKQTNTDKEKIKE
KGSFSDTGLGDGKMKSDSFAPKTDSEKPFRG
SQSPKRYKLRDDFEKKMADFHKEEMDDQDK
DKAKGRKESEFDDEPKFMSKVIGANKNQEEB
KSGKWEGLVYAPPGKEKQRKTEELEEESFPE
RSKKEDRGKRSEGGHRGFVPEKNFRVTAYK
AVQEKSSSPPPRKTSESRDKLGAKGDFPTGKS
SFSITREAQVNVRMDSFDEDLARPSGLLAQER
KLCRDLVHSNKKEQEFRSIFQHIQSAQSQRSP
SELFAQHIVTIVHHVKEHHFGSSGMTLHERFT
KYLKRGTEQEAAKNKKSPEIHRRIDISPSTFRK
HGLAHDEMKSPREPGYKAEGKYKDDPVDLR
LDIERRKKHKERDLKRGKSRESVDSRDSSHSR
ERSAEKTEKTHKGSKKQKKHRRARDRSRSSS
SSSQSSHSYKAEEYTEETEEREESTTGFDKSRI

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion GTKDFVGPSERGGGRARGTFQFRARGRGWGRGNYSGNNNNNSNNDFQKRNREEEWDPEYT
997	2347		8209	202		PKSKKYYLHDDREGEGSDKWVSRGRGRGAF PRGRGRFMFRKSSTSPKWAHDKFSGEEGEIE DDESGTENREEKDNIQPTTE
		A	8398	202	552	CPALGGRQDLQGTRLLWAHDSGVGGQKAKS KQENLESLEATGREEEGGQGPPVTTKGVLLA LLMAGLALQPGTALLCYSCKAQVSNEDCLQ VENCTQLGEQCWTARIREWGDDSRQA
998	2348	A	8400	697	301	NPPSACTPGSCDSCSGRGRDLAFDSVWSTNN MSDPRRPNKVLRYKPPPSECNPALDDPTPDY MNLLGMIFSMCGLMLKLKWCAWVAVYCSFI SFANSRSSEDTKQMMSSFMLSISAVVMSYLQ NPQPMTPPW
999	2349	A	8401	93	1126	ASASHITSGHLRCFPGSEGVGTMARCFSLVLL LTSIWTTRLLVQGSLRAEELSIQVSCRIMGITL VSKKANQQLNFTEAKEACRLLGLSLAGKDQ VETALKASFETCSYGWVGDGFVVISRISPNPK CGKNGVGVLIWKVPVSRQFAAYCYNSSDTW TNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYS VASPYSTIPAPTTTPPAPASTSIPRRKKLICVTE VFMETSTMSTETEPFVENKAAGLGFCYVKRYVKAF PFTNKNQQKEMIETKVVKEEKANDSNPNEES KKTDKNPEESKSPSKTTMRCLEAEV
1000	2350	A	8406	2	777	KERCQFVVKPMLSTVGSFLQDLQNEDKGIKT AAIFTADGNMISASTLMDILLMNDFKLVINKI AYDVQCPKREKPSNEHTAEMEHMKSLVHRL FTILHLEESQKKREHHLLEKIDHLKEQLQPLE QVKAGIEAHSEAKTSGLLWAGLALLSIQGGA LAWLTWWYSWDIMEPVTYJTFANSMVFF AYFIVTRQDYTYSAVKSRQFLQFFHKKSKQQ HFDVQQYNKLKEDLAKAKESLKQARHSLCL QMQVEELNEKN
1001	2351	A	8410	1400	264	VGFWERPLRSSRWFRRSLRRWEMLARAARG TGALLLRGSLLASGRAPRRASSGLPRNTVVLF VPQQEAWVVERMGRFHRILEPGLNILIPVLDR IRYVQSLKEIVINVPEQSAVTLDNVTLQIDGV LYLRIMDPYKASYGVEDPEYAVTQLAQTTM RSELGKLSLDKVFRERESLNASIVDAINQAAD CWGIRCLRYEIKDIHVPPRVKESMQMQVEAE RRKRATVLESEGTRESAINVAEGKKQAQILAS EAEKAEQINQAAGEASAVLAKAKAKAEIRI LAAALTQHNGDAAASLTVAEQYVSAFSKLA KDSNTILLPSNPGDVTSMVAQAMGVYGALT KAPVPGTPDSLSSGSSRDVQGTDASLDEELDR VKMS
	2352	A	8421	134	941	NRENLLESRMMDPCSVGVQLRTTNECHKTY YTRHTGFKTLQELSSNDMLLLQLRTGMTLSG NNTICFHHVKIYIDRFEDLQKSCCDPFNIHKKL AKKNLHVIDLDDATFLSAKFGRQLVPGWKLC PKCTQIINGSVDVDTEDRQKKKPESDGRTAK ALRSLQFTNPGRQTEFAPETGKREKRRLTKN ATAGSDRQVIPAKSKVYDSQGLLIFSGMDLC DCLDEDCLGCFYACPACGSTKCGAECRCDRK WLYEQIEIEGGEIHNKHAG
1003	2353	A	8427	3	1416	TEWGLSGSCPGCSPLEPGSRGRGAAAWRILR CRRLPEPSPFLTQPNLAQSQPPAPVPVTDPSVT MHPAVFLSLPDLRCSLLLLVTWVFTPVTTEIT

SEQ ID	SEQ ID	Met	Teen	Dandistad	Dunding - 1 1	L Amino cold commune (AmAlanian C. C
			SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ĺ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
-psz	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	[1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
İ	1	ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
ì	j	1		peptide		/=possible nucleotide deletion, \=possible
1		ì		sequence		nucleotide insertion
		 	 	sequence		
ł	1		ł		ł	SLDTENIDEILNNADVALVNFYADWCRFSQM
l		ľ			,	LHPIFEEASDVIKEEFPNENQVVFARVDCDQH
1						SDIAQRYRISKYPTLKLFRNGMMMKREYRGQ
1	1					RSVKALADYIRQQKSDPIQEIRDLAEITTLDRS
1	ł	l		l		KRNIIGYFEQKDSDNYRVFERVANILHDDCAF
i	1	Ĭ	1	i	{	LSAFGDVSKPERYSGDNIIYKPPGHSAPDMVY
1			1			LGAMTNFDVTYNWIQDKCVPLVREITFENGE
		1	1		i	ELTEEGLPFLILFIIMKEDTESLEIFQNEVARQL
i	i	i		1	f	ISEKGTINFLHADCDKFRHPLLHIOKTPADCP
1		l		i		1
		1			1	VIAIDSFRHMYVFGDFKDVLIPGKLKQFVFDI.
1	i	i	1	ĺ	i	HSGKLHREFHHGPDPTDTAPGEQAQDVASSP
		<u> </u>				PESSFQKLAPSEYRYTLLRDRDEL
1004	2354	Α	8432	910	387	GLSRKLRAGFLPGFCRVSPCGSWVVETLVKM
l	1	Į .	į į			ACAAARSPADQDRFICIYPAYLNNKKTIAEGR
i		1				RIPISKAVENPTATEIQDVCSAVGLNVFLEKN
]					•	KMYSREWNRDVQYRGRVRVQLKQEDGSLC
1	ļ	l				LVQFPSRKSVMLYAAEMIPKLKTRTQKTGGA
1	1					
1005	2355	A	8453	90	530	DQSLQQGEGSKKGKGKKKKK
1,002	2333	^	0433	90	330	QSHETKMQSGTHWRVLGLCLLSVGVWGQD
	1					GNEEMGGITQTPYKVSISGTTVILTCPQYPGSE
i .	·	ł	1			ILWQHNDKNIGGDEDDKNIGSDEDHLSLKEF
l		ĺ				SELEQSGYYVCYPRGSKPEDANFYLYLRARG
L			1			NPGLQNRYHRLFREDHSKGHSQ
1006	2356	Α	8458	3	307	AVQRIRHEMNIFRLTGDLSHLAAIVILLLKIW
i '	1			·	,	KTRSCAGISGKSQLLFALVFTTRYLDLFTSFIS
						LYNTSMKVWYAIHRNVFHLQCTGLWTLNLC
1	l			•		QLCIFN TAILING THE QUICK TENDE
1007	2357	A	8459	43	553	
1007	2337	^	0439	43	333	GAGAGGDWAAMDKLKKVLSGQDTEDRSGL.
1	ļ	J				SEVVEASSLSWSTRIKGFIACFAIGILCSLLGT
	ĺ					VLLWVPRKGLHLFAVFYTFGNIASIGSTIFLM
1		ļ				GPVKQLKRMFEPTRLIATIMVLLCFALTLCSA
	ļ	l))			FWWHNKGLALIFCILQSLALTWYSLSFIPFAR
						DAVKKCFAVCLA
1008	2358	A	8462	487	150	AQDIRSVHSLGQKSTFVKHFRTLSHLHGLPDP
						PPHWPPQERSPPSHPCMPSHRPQIPQLSNSGPS
] [ĺ		DPRWGCVGPSMPTSTCLPGAVEASTTKASLP
				l	1	KCPVDSSLPTPEACFL
1009	2359	A	8465	134	054	
1007	23.27	^	0403	134	954	ETRVKTSLELLRTQLEPTGTVGNTIMTSQPVP
				j	,	NETTIVLPSNVINFSQAEKPEPTNQGQDSLKKH
					l	LHAEIKVIGTIQILCGMMVLSLGIILASASFSPN
				l	l	FTQVTSTLLNSAYPFIGPFFFIISGSLSIATEKRL
				j	· J	TKLLVHSSLVGSILSALSALVGFIILSVKQATL
						NPASLQCELDKNNIPTRSYVSYFYHDSLYTTD
						CYTAKASLAGTLSLMLICTLLEFCLAVLTAVL
				J	ļ	RWKQAYSDFPGSVLFLPHSYIGNSGMSSKMT
						HDCGYEELLTS
1010	2360	A	8468	2	473	
****	2000	^	0400	-	713	KYRYRRPYPVMRKICQVGPAGLAFILNISPVA
				1		HRVALCHLAGCQEQAAWYHTLQILFFLVSAY
	•					FFSCPVPEKYFPGSCDIVGHGHQIFHAFLSICI
						LSQLEAILLDYQGRQEIFLQRHGPLSVHMACL
				f		SFFFLAACSAATAALLRHKVKARLTKKDS
1011	2361	Α	8478	5	409	TELSQLEKAHPPADMGRRKSKRKPPPKKKMT
			[GTLETQFTCPFCNHEKSCDVKMDRARNTGVI
]		SCTVCLEEFQTPITCILGNLGFFQRVGRGLESG
				}	j	
				1	1	PCSSGPLCALVQGQSRPEEQVPPSDFCGVRRC
1012	2262	\vdash	0401	2010	1750	RAGFQCQ
1012	2362	A	8481	2810	1652	RTSTQKWQSVFNDSQEHLERFYCNPENDRM
			l	ļ		RMKYGGQEFWADLNAMNVYETTEFDQLRR
L						LSTPPSSNVNSIYHTVWKFFCRDHFGWREYPE
						·····

SEQ ID NO: of NO: of NO: of nucleotide peptide cotide sequence USSN 09/496 USSN	LHINS TQAP YMHP EFKYR MFGR PRVCG SKGV NPGS PYFVI
nucl- cotide seq- uence Destide Destide seq- uence Des	LHINS TQAP YMHP EFKYR MFGR RVCG SKGV NPGS PYFVI EFSCL LRFTP
Corresponding to last amino acid residue of peptide sequence S	LHINS TQAP YMHP EFKYR MFGR RVCG SKGV NPGS PYFVI EFSCL LRFTP
sequence Sequence Sequence O9/496 Orderspondi ng to first amino acid residue of peptide sequence Sequence Threonine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide de	TQAP YMHP EFKYR MFGR PRVCG SKGV NPGS PYFVI GFSCL CRFTP
uence 914 ng to first amino acid residue of peptide sequence T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, Y=Tyrosine, X=Unknown, *=Stop codon, Y=Tyrosine, X=Unknown, *=Stop codon, Y=Tyrosine nucleotide deletion, V=Possible nucleotide insertion	TQAP YMHP EFKYR MFGR PRVCG SKGV NPGS PYFVI GFSCL CRFTP
amino acid residue of peptide sequence	TQAP YMHP EFKYR MFGR PRVCG SKGV NPGS PYFVI GFSCL CRFTP
residue of peptide sequence	TQAP YMHP EFKYR MFGR PRVCG SKGV NPGS PYFVI GFSCL CRFTP
Popsible nucleotide deletion, \=possible nucleotide insertion SVIRLIEEANSRGI.KEVRFMMWNNHY FFREIKRRPLFRSCFILLPYLQTLGGVF PPLEATSSQUICPDGVTSANFYPETWV SQDFIQVPVSAEDKSYRIIYNLFHKTVP ILQILRVQNQFLWEKYKRKKEYMNRK DRIINREHLPHGTSQDVVDGICKHNFDI KHATMFGQGSYFAKKASYSHNFSKKS: HFMFLAKVLTGRYTMGSHGMRRPPV VTSDLYDSCVDNFFEPQIFVIFNDDQSY QYEEVSNTVSI IENCRTRLRQAWHEVCGNKMAAPIPQC SRFLGWWFRQPVLVTQSAAIVPVRTKK PIYQPKFKTEKEFMQHARKAGLVIPPEK IHLACTAGIFDAYVPPEGDARISSLSKE TERMKKTMASQVSIRRIKDYDANFKIK KAKDIFIEGSPLY IRTGYVYLTHSLCIHMYSIRTAYVY AQLMYTYVFYTHRLCIHMYSIRTAYVY AQLMYTYVFYTHRLCHIMYSIRTAYVY AQLMYTYVFYTHRLCHIMYSIRTDYVY AQLMYTYVFYTHRLCHIMYSIRTDYVY AQLMYTYVFYTHSYMSDE INSSEHFSQAPQRLSFYSWYGSARLFRFF AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLST NASVNVSHPAPGDWFVAAHLPPSSQKI LAPTCAYVFQPELLVTRVVEISIMEPDV.	TQAP YMHP EFKYR MFGR PRVCG SKGV NPGS PYFVI GFSCL CRFTP
Sequence nucleotide insertion	TQAP YMHP EFKYR MFGR PRVCG SKGV NPGS PYFVI GFSCL CRFTP
SVIRLIEEANSRGLKEVRFMMWNNHYI FFRREIKRRPLFRSCFILLPYLQTLGGVF PPLEATSSSQIICPDGVTSANFYPETWV SQDFIQVPVSAEDKSYRIIYNLFHKTVP ILQILRVQNQFLWEKYKRKEYMNRK DRIINERHLPHGTSQDVVDGICKHNFDI KHATMFGQGSYFAKKASYSHNFSKKS HFMFLAKVLTGRYTMGSHGMRRPPPV VTSDLYDSCVDNFFEPQIFVIFNDDQSY QYEEVSNTVSI 1013 2363 A 8488 2 517 IENCRTRLRQAWHEVCGNKMAAPIPQC SRFLGWWFRQPVLVTQSAAIVPVRTKK PIYQPKFKTEKEFMQHARKAGLVIPPEK IHLACTAGIFDAYVPPEGDARISSLSKEC TERMKKTMASQVSIRRIKDYDANFKIK KAKDIFIEGSPLY 1014 2364 A 8501 363 17 YIRTGYVYICIIYAQLMYTYYIRTAYVY AQLMYTYVLYTHSLCHMYSIRTAYVY AQLMYTYVLYTHSLCHMYSIRTAYVY AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFF PVINPLGTSFPDDTAVQPSFQVGVPLST NASVNVSHPAPGDWFVAAHLPPSSQKII LAPTCAYVFQPELLVTRVVEISIMEPDV	TQAP YMHP EFKYR MFGR PRVCG SKGV NPGS PYFVI GFSCL CRFTP
FFRREIKRRPLFRSCFILLPYLQTLGGVF PPLEATSSSQIICPDGVTSANFYPETWV SQDFIQVPVSAEDKSYRIIYNLFHKTVP ILQILRVQNQFLWEKYKRKKEYMNRK DRIINERHLPHGTSQDVVDGICKHNFDH KHATMFGQGSYFAKKASYSHNFSKKS HFMFLAKVLTGRYTMGSHGMRRPPPV VTSDLYDSCVDNFFEPQIFVIFNDDQSY QYEEVSNTVSI 1013 2363 A 8488 2 517 IENCRTRLRQAWHEVCGNKMAAPIPQC SRFLGWWFRQPVLVTQSAAIVPVRTKK PIYQPKFKTEKEFMQHARKAGLVIPPEK IHLACTAGIFDA YVPPEGDARISSLSKEC TERMKKTMASQVSIRRIKDYDANFKIK KAKDIFIEGSPLY AQLMYTYVLYTHSLCIHMYSIRTAYVY AQLMYTYVFYTHRLCIHMYSIRTAYVY AQLMYTYVFYTHRLCIHMYSIRTDYVY AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFF AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLS NASVNVSHPAPGDWFVAAHLPPSSQKII LAPTCAYVFQPELLVTRVVEISIMEPDV	TQAP YMHP EFKYR MFGR PRVCG SKGV NPGS PYFVI GFSCL CRFTP
PPLEATSSSQIICPDGVTSANFYPETWV SQDFIQVPVSAEDKSYRIIYNLFHKTVP ILQILRVQNQFLWEKYKKKEYMNRK DRIINERHLFHGTSQDVVDGICKHNFDI KHATMFGQGSYFAKKASYSHNFSKKS HFMFLAKVLTGRYTMGSHGMRRPPPV VTSDLYDSCVDNFFEPQIFVIFNDDQSY QYEEVSNTVSI 1013 2363 A 8488 2 517 IENCRTRLRQAWHEVCGNKMAAPIPQC SRFLGWWFRQPVLVTQSAAIVPVRTKK PIYQPKFKTEKEFMQHARKAGLVIPPEK IHLACTAGIFDAYVPPEGDARISSLSKEC TERMKKTMASQVSIRRIKDYDANFKIK KAKDIFIEGSPLY 1014 2364 A 8501 363 17 YIRTGYVYICIIYAQLMYTYYIRTAYVY AQLMYTYVFYTHRLCIHMYSIRTDYVY AQLMYTTYVFYTHSLCIHMYSIRTDYVY AQLMYTTYFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFR AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLS NASVNVSHPAPGDWFVAAHLPPSSQKII LAPTCAYVFQPELLVTRVVEISIMEPDV	YMHP EFKYR MFGR PRVCG SKGV NPGS PYFVI FFSCL ERFTP
SQDFIQVPVSAEDKSYRIIYNLFHKTVP. II.QILRVQNQFL WEKYKRKEYMNRK. DRIINERHLPHGTSQDVVDGICKHNFDI KHATMFGQGSYFAKKASYSHNFSKKS. HFMFLAKVLTGRYTMGSHGMRRPPPV VTSDLYDSCVDNFFEPQIFVIFNDDQSY QYEEVSNTVSI 1013 2363 A 8488 2 517 IENCRTRLRQAWHEVCGNKMAAPIPQC SRFLGWWFRQPVLVTQSAAIVPVRTKK PIYQPKFKTEKEFMQHARKAGLVIPPEK IHLACTAGIFDAYVPPEGDARISSLSKEC TERMKKTMASQVSIRRIKDYDANFKIK KAKDIFIEGSPLY AQLMYTYVLYTHSLCHMYSIRTAYVY AQLMYTYVFYTHRLCHMYSIRTAYVY AQLMYTYVFYTHRLCHMYSIRTDYVY AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFR AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLS. NASVNVSHPAPGDWFVAAHLPPSSQKII LAPTCAYVFQPELLVTRVVEISIMEPDV	EFKYR MFGR PRVCG SKGV NPGS PYFVI FFSCL IRFTP
ILQILRVQNQFLWEKYKRKKEYMNRK. DRIINERHLPHGTSQDVVDGICKHNFDI KHATMFGQGSYFAKKASYSHNFSKKS. HFMFLAKVLTGRYTMGSHGMRRPPPV VTSDLYDSCVDNFFEPQIFVIFNDDQSY QYEEVSNTVSI 1013 2363 A 8488 2 517 IENCRTRLRQAWHEVCGNKMAAPIPQC SRFLGWWFRQPVLVTQSAAIVPVRTKK PIYQPKFKTEKEFMQHARKAGLVIPPEK IHLACTAGIFDA YVPPEGDARISSLSKEC TERMKKTMASQVSIRRIKDYDANFKIK KAKDIFIEGSPLY AQLMYTYVLYTHSLCIHMYSIRTAYVY AQLMYTYVLYTHSLCIHMYSIRTAYVY AQLMYTYVFYTHRLCIHMYSIRTDYVY AQLMYTYVFYTHRLCIHMYSIRTDYVY AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFF AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLS: NASVNVSHPAPGDWFVAAHLPPSSQKII LAPTCAYVFQPELLVTRVVEISIMEPDV	MFGR PRVCG SKGV NPGS PYFVI FFSCL RFTP
DRIINERHLPHGTSQDVVDGICKHNFDI KHATMFQQGSYFAKKASYSHNFSKKS. HFMFLAKVLTGRYTMGSHGMRRPPPV VTSDLYDSCVDNFFEPQIFVIFNDDQSY QYEEVSNTVSI 1013 2363 A 8488 2 517 IENCRTRLRQAWHEVCGNKMAAPIPQC SRFLGWWFRQPVLVTQSAAIVPVRTKK PIYQPKFKTEKEFMQHARKAGLVIPPEK IHLACTAGIFDAYVPPEGDARISSLSKE TERMKTMASQVSIRRIKDYDANFKIK KAKDIFIEGSPLY 1014 2364 A 8501 363 17 YIRTGYVYICIIYAQLMYTYYIRTAYVY AQLMYTYVLYTHSLCIHMYSIRTDYVY. AQLMYTYVFYTHRLCIHMYSIRTDYVY. AQLMYTYVFYTHSVMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFR AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLS: NASYNVSHPAPGDWFVAAHLPPSSQKII LAPTCAYVFQPELLVTRVVEISIMEPDV.	PRVCG SKGV NPGS PYFVI GFSCL CRFTP
Control of the cont	SKGV NPGS PYFVI SFSCL CRFTP
HFMFLAKVLTGRYTMGSHGMRRPPPV VTSDLYDSCVDNFFEPQIFVIFNDDQSY QYEEVSNTVSI	NPGS PYFVI FSCL RFTP
VTSDLYDSCVDNFFEPQIFVIFNDDQSY QYEEVSNTVSI	PYFVI FSCL RFTP
1013 2363 A 8488 2 517 IENCRTRLRQAWHEVCGNKMAAPIPQC SRFLGWWFRQPVLVTQSAAIVPVRTKK PIYQPKFKTEKEFMQHARKAGLVIPPEK IHLACTAGIFDA YVPPEGDARISSLSKEC TERMKKTMASQVSIRRIKDYDANFKIK KAKDIFIEGSPLY 1014 2364 A 8501 363 17 YIRTGYVYICIIYAQLMYTYYIRTAYVY AQLMYTYVFYTHSLCIHMYSIRTDYVY AQLMYTYVFYTHSLCIHMYSIRTDYVY AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFF AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLS: NASVNVSHPAPGDWFVAAHLPPSSQKII LAPTCAYVFQPELLVTRVVEISIMEPDV	FSCL RFTP
1013 2363 A 8488 2 517 IENCRTRLRQAWHEVCGNKMAAPIPQC SRFLGWWFRQPVLVTQSAAIVPVRTKK PIYQPKFKTEKEFMQHARKAGLVIPPER IHLACTAGIFDAYVPPEGDARISSLSKEG TERMKKTMASQVSIRRIKDYDANFKIK KAKDIFIEGSPLY 1014 2364 A 8501 363 17 YIRTGYVYICIIYAQLMYTYYIRTAYVY AQLMYTYVLYTHSLCIHMYSIRTAYVY AQLMYTYVFYTHRLCIHMYSIRTDYVY AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFR AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLS NASVNVSHPAPGDWFVAAHLPPSSQKII LAPTCAYVFQPELLVTRVVEISIMEPDV	RFTP
SRFLGWWFRQPVLVTQSAAIVPVRTKK PIYQPKFKTEKEFMQHARKAGLVIPPER IHLACTAGIFDAYVPPEGDARISSLSKEG TERMKKTMASQVSIRRIKDYDANFKIK KAKDIFIEGSPLY 1014 2364 A 8501 363 17 YIRTGYVYICIIYAQLMYTYYIRTAYVY AQLMYTYVLYTHSLCIHMYSIRTAYVY AQLMYTYVFYTHRLCIHMYSIRTDYVY AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFR AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLS NASVNVSHPAPGDWFVAAHLPPSSQKI LAPTCAYVFQPELLVTRVVEISIMEPDV	RFTP
PIYQPKFKTEKEFMQHARKAGLVIPPEK IHLACTAGIFDAYVPPEGDARISSLSKEG TERMKKTMASQVSIRRIKDYDANFKIK KAKDIFIEGSPLY 1014 2364 A 8501 363 17 YIRTGYVYICIIYAQLMYTYYIRTAYVY AQLMYTYVLYTHSLCIHMYSIRTAYVY AQLMYTYVFYTHRLCIHMYSIRTDYVY AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFR AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLSI NASVNVSHPAPGDWFVAAHLPPSSQKI LAPTCAYVFQPELLVTRVVEISIMEPDV	
IHLACTAGIFDAYVPPEGDARISSLSKEG TERMKKTMASQVSIRRIKDYDANFKIK KAKDIFIEGSPLY 1014 2364 A 8501 363 17 YIRTGYVYICIIYAQLMYTYYIRTAYVY AQLMYTYVLYTHSLCIHMYSIRTAYVY AQLMYTYVFYTHRLCIHMYSIRTDYVY. AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFF AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLSI NASVNVSHPAPGDWFVAAHLPPSSQKI LAPTCAYVFQPELLVTRVVEISIMEPDV	SDRS
1014 2364 A 8501 363 17 YIRTGYVYICIIYAQLMYTYYIRTAYVY AQLMYTYVLYTHSLCIHMYSIRTAYVY AQLMYTYVFYTHRLCIHMYSIRTDYVY AQLMYTYVFYTHRLCIHMYSIRTDYVY AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFF AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLS: NASVNVSHPAPGDWFVAAHLPPSSQKI LAPTCAYVFQPELLVTRVVEISIMEPDV	
1014 2364 A 8501 363 17 YIRTGYVYICIIYAQLMYTYYIRTAYVY AQLMYTYVLYTHSLCIHMYSIRTAYVY AQLMYTYVFYTHRLCIHMYSIRTDYVY AQLMYTYVFYTHRLCIHMYSIRTDYVY AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFF AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLS: NASVNVSHPAPGDWFVAAHLPPSSQKI LAPTCAYVFQPELLVTRVVEISIMEPDV	
1014 2364 A 8501 363 17 YIRTGYVYICIIYAQLMYTYYIRTAYVY AQLMYTYVFYTHSLCIHMYSIRTAYVY AQLMYTYVFYTHRLCIHMYSIRTDYVY AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFF AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLS: NASVNVSHPAPGDWFVAAHLPPSSQKII LAPTCAYVFQPELLVTRVVEISIMEPDV	DFPE
AQLMYTYVLYTHSLCIHMYSIRTAYVY AQIMYTYVFYTHRLCIHMYSIRTDYVY AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFF AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLSI NASVNVSHPAPGDWFVAAHLPPSSQKI LAPTCAYVFQPELLVTRVVEISIMEPDV	
AQIMYTYVFYTHRLCIHMYSIRTDYVY. AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFF AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLST NASVNVSHPAPGDWFVAAHLPPSSQKIL LAPTCAYVFQPELLVTRVVEISIMEPDV	
AQLMYTYVFYTHSYMSDE 1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFF AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLST NASVNVSHPAPGDWFVAAHLPPSSQKIL LAPTCAYVFQPELLVTRVVEISIMEPDV	
1015 2365 A 8504 3 2190 NSSEHFSQAPQRLSFYSWYGSARLFRFF AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLST NASVNVSHPAPGDWFVAAHLPPSSQKIL LAPTCAYVFQPELLVTRVVEISIMEPDV	CILY
AVLLRWLLQVSRESGAACTDAEITVHF PPVINPLGTSFPDDTAVQPSFQVGVPLS NASVNVSHPAPGDWFVAAHLPPSSQKI LAPTCAYVFQPELLVTRVVEISIMEPDV	
PPVINPLGTSFPDDTAVQPSFQVGVPLST NASVNVSHPAPGDWFVAAHLPPSSQKI LAPTCAYVFQPELLVTRVVEISIMEPDV	VPPD
NASVNVSHPAPGDWFVAAHLPPSSQKI LAPTCAYVFQPELLVTRVVEISIMEPDV	RSGA
LAPTCAYVFQPELLVTRVVEISIMEPDV	
	ELKG
	PLPQ
TLLSHPSYLKVFVPDYTRELLLELRDCV	SNGS
LGCPVRLTVGPVTLPSNFQKVLTCTGAI	'WPC
RLLLPSPPWDRWLQVTAESLVGPLGTV	AFSA
VAALTACRPRSVTIQPLLQSSQNQSFNA	SSGL
LSPSPDHQDLGRSGRVDRSPFCLTNYPV	TRED
MDVVSVHFQPLDRVSVRVCSDTPSVMF	
NTGMDSGGSLTISLRANKTEMRNETVV	VACV
NAASPFLGFNTSLNCTTAFFQGYPLSLS.	
RANLIIPYPETDNWYLSLQLMCPENAED	
AVVHVETTLYLVPCLNDCGPYGOCLLL	
YLYASCSCKAGWRGWSCTDNSTAQTV	
. AATLLLTLSNLMFLAPIAVSVRRFFLVE	
AYTMFFSTFYHACDQPGEAVLCILSYD1	
CDFLGSGAAIWVTILCMARLKTVLKYV	
GTLVIAMSLQLDRRGMWNMLGPCLFAI	
ASMWAYRCGHRRQCYPTSWQRWAFYI	
VSMASVGIAIYTSMMTSDNYYYTHSIW	an i
AGSAALLLPPPDQPAEPWACSQKFPCHY	OTC
KNDREELYAVT	٧.٠
1016 2366 A 8511 1 453 KWYPSOPVRIPGRFYYKLPAGHRRCRM	ADAV
KGGEKKKGRSAINEVVTREYTINIHKRIH	
FKKRAPRALKEIRKFAMKEMGTPDVRII	
NKAVWAKGIRNVPYRIRVRLSRKRNED	
	3D2b
1017 2367 A 8513 54 1196 LERTPASADMAWTK VOLET AGT MI VTC	
TENTITATION TO THE TOTAL A LET OF THE LAND THE A LET OF THE LAND THE A LET OF THE LAND THE LAND THE A LET OF THE LAND THE	TAICI
LSAKWADNFMAEGCGGSKEHSFQHPFL	UAV
GMFLGEFSCLAAFYLLRCRAAGQSDSSN	
QPFNPLLFLPPALCDMTGTSLMYVALNN	
SSFQMLRGAVIIFTGLFSVAFLGRRLVLS	
GILATIAGLVVVGLADLLSKHDSQHKLS	QWL
GDLLIIMAQIIVAIQMVLEEKFVYKHNVI	QWL EVIT
AVGTEGLFGFVILSLLLVPMYYIPAGSFS	QWL EVIT IPLR
RGTLEDALDAFCQVGQQPLIAVALLGNI	QWL EVIT IPLR GNP
FFNFAGISVTKELSATTRMVLDSLRTVVI	QWL EVIT IPLR GNP SSIA

DEC ID	SEO ID	Met	SEQ	Predicted	Dendict	Aming gold sequence (A=Alesias C=Costs
SEQ ID	SEQ ID	hod	ID NO:		Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	1100		beginning		
nucl-	peptide		in	nucleotide	location	F-Phenylalanine, G-Glycine, H-Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	j	l	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
}	1	ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		i		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	Į.			peptide		/=possible nucleotide deletion, \=possible
1				sequence	•	nucleotide insertion
						SLALGWEAFHALQILGFLILLIGTALYNGLHR
1	ł	1		ļ		PLLGRLSRGRPLAEESEQERLLGGTRTPINDA
j				ĺ		S
1010	22.50	ļ. <u>.</u>	0510	224	604	
1018	2368	A	8518	324	694	SPFWTEKRRMEKPLFPLVPLHWFGFGYTALV
ļ		ļ	l		•	VSGGIVGYVKTGSVPSLAAGLLFGSLAGLGA
1	İ	1		Ì		YQLYQDPRNVWGFLAATSVTFVGVMGMRS
İ	1	ľ				YYYGKFMPVGLIAGASLLMAAKVGVRMLM
	ŧ					TSD
1019	2369	A	8526	2	1787	VSAAAVNMEPPDAPAQARGAPRLLLLAVLL
ł	!	1	i	ļ		AAHPDAQAEVRLSVPPLVEVMRGKSVILDCT
1	ł	ł	ł	i	1	PTGTHDHYMLEWFLTDRSGARPRLASAEMO
	ŀ		İ		ì	GSELQVTMHDTRGRSPPYQLDSQGRLVLAEA
	ĺ	ļ			•	OVGDERDYVCVVRAGAAGTAEAAARLNVF
1	ĺ	ĺ				`
]	1	<u> </u>			AKPEATEVSPNKGTLSVMEDSAQEIATSNSRN
	1	1			Į	GNPAPKITWYRNGQRLEVPVEMNPEGYMTS
					ì	RTVREASGLLSLTSTLYLRLRKDDRDASFHC
Į	ļ	l	ļ	Į.	1	AAHYSLPEGRHGRLDSPTFHLTLHYPTEHVQ
İ	1					FWVGSPSTPAGWVREGDTVQLLCRGDGSPSP
Į	l					EYTLFRLQDEQEEVLNVNLEGNLTLEGVTRG
!	ł	ł	[1		QSGTYGCRVEDYDAADDVQLSKTLELRVAY
1			1			LDPLELSEGKVLSLPLNSRAVVNCSVHGLPTP
[l]			ALRWTKDSTPLGDGPMLSLSSITFDSNGTYVC
1						EASLPTVPVLSRTQNFTLLVQGSPELKTAEIEP
l	ļ	J	}	J		KADGSWREGDEVTLICSARGHPDPKLSWSQL
İ	İ		1			GGSPAEPIPGRQGWVSSSLTLKVTSALSRDGI
1	1					SCEASNPHGNKRHVFHFGTVSPQTSQAGVAV
ł	Ì	ļ		,		MAVAVSVGLLLLVVAVFYCVRRKGGPCCRO
1				•		
1000	2270	<u> </u>	0600	2	1200	RREKGAP
1020	2370	A	8530	Z.	1200	PRVRLLRPSRSRSCRGLLSTRAPGPSPFRSLHS
]		ļ	SPLLPHAMKSPFYRCQNTTSVEKGNSAVMGG
	ł	1				VLFSTGLLGNLLALGLLARSGLGWCSRRPLR
i	i	1				PLPSVFYMLVCGLTVTDLLGKCLLSPVVLAA
ļ	ļ	1			ļ	YAQNRSLRVLAPALDNSLCQAFAFFMSFFGL
						SSTLQLLAMALECWLSLGHPFFYRRHITLRLG
ł						ALVAPVVSAFSLAFCALPFMGFGKFVQYCPG
1	!	1				TWCFIQMVHEEGSLSVLGYSVLYSSLMALLV
]	}	J	J		J	LATVLCNLGAMRNLYAMHRRLORHPRSCTR
1						DCAEPRADGREASPQPLEELDHLLLLALMTV
1		ł			[LFTMCSLPVIYRAYYGAFKDVKEKNRTSEEA
	l	!				EDLRALRFLSVISIVDPWIFIIFRSPVFRIFFHKI
1					!	
1021	2271	 	9526		127	PROFILMATECOVEL ELETETS CREPORENDO
1021	2371	A	8536	1	237	RRGEIDMATEGDVELELETETSGPERPPEKPR
	l .	1	1			KHDSGAADLERVTDYAEEKEIQSSNLETAMS
	<u> </u>		L			VIGDRRSREQKAKQER
1022	2372	A	8537	94	541	RKERRRRRRMEAVVFVFSLLDCCALIFLSV
1]	ļ			Į	YFITTLSDLECDYINARSCCSKLNKWVIPELIG
	I					HTIVTVLLLMSLHWFIFLLNLPVATWNIYRYI
1	1	ļ			ł	MVPSGNMGVFDPTEIHNRGQLKSHMKEAMI
1		ĺ	İ		·	KLGFHLLCFFMYLYSMILALIND
1023	2373	A	8540	26	431	RMMKCPQALLAIFWLLLSWVSSEDKVVQSPL
1.0~	~	l "	0540	~0	7,5	SLVVHEGDTVTLNCSYEVTNFRSLLWYKQEK
1	1					LA MERI DA TOUCIERRO DE DER DARES COM
	1	1		i		KAPTFLFMLTSSGIEKKSGRLSSILDKKELSSIL
1	ĺ	1				NITATQTGDSAIYLCAVEAQCSLVTCSLYSNS
		L			<u></u> _	TAEALQL
1024	2374	À	8544	1731	743	GVRLRYSPIAVVMVGEAGRDLRRRRAVAVT
	l	1	ľ		Ī	AEKMAVLAPLIALVYSVPRLSRWLAQPYYLL
	Į	ł	1	1	ł	SALLSAAFLLVRKLPPLCHGLPTQREDGNPCD
	l	1			l	FDWREVEILMFLSAIVMMKNRRSITVEQHIGN
	1	1			1	IFMFSKVANTILFFRLDIRMGLLYITLCIVFLM
		L		L _	1	

NO: of No: of N	SEQ ID	SEQ ID	Met	SEQ	Predicted	Deadisted	Amino poid spane (A-Alexia - C. C.
nuciocide sequence where the corresponding of the c				, ,		Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
Solid Sequence O.			l non	1 .			
Seq		1					
Benne			1			to lest emino	M=Methionine N=Asperagine D=Proline
mino acid residue of poptide sequence se		40					
residue of poptible sequence	3050	l	l	1			
ppptide				1			
	1		ł				
1025			į	}			nucleotide insertion
WIVEFFANWSINDCOSFAPIYADLSILKYNCTG	-	 	 	 	3-1		TCKPPLYMGPEYIKYFNDKTIDEFI ERDKRYT
I.NFGKVDVGRYTDVSTRYKVSTSPLTKQLPT	1			-			
LLEPGGGKEAMRRPQIDKKGRAVSWTFSEEN							LNFGKVDVGRYTDVSTRYKVSTSPLTKOLPT
1025 2375 A 8546 2194 1707 TYSPIENINELYQRAKKLSKAGDNIPEEQPVAS TPITYSDOENKKOK TYSPIKITMASI.KCSTVVCVICLEKEY KIRCS TYSPIKITMASI.KCSTVVCVICLEKEY KIRCS CRYPVCSVVCTREKKEQOMETRYVEKIRS ALPTKTVKPVENKDDDDSIADFUNSDEEDR VSLONLKNLGESATLKSILLINPHLRQLKVVI DQGEDKAKLMRAYMGEPLFVERADCCLGIV EPSQNESS ALPTKTVKPVENKDDDDSIADFUNSDEEDR VSLONLKNLGESATLKSILLINPHLRQLKVI DQGEDKAKLMRAYMGEPLFVERADCCLGIV EPSQNESS VGMEIPAVNILKVILLGHWILLTTWGCIVFSGS YAWANFILALGWWAVAQRDSIDAISMFLGG LLATTHEDIVISISIPYRVSILDTORRGVGMAIL SILLKPLSCCFVYHMYRERGGELLVHTGFG SQDRSAYQTIDSAEAADFAPAVPFAVPGGRSQDAR GY DPI-GPASPQEEGGSESSTMTELETAMGMIIDV FSRYSGSEGSTQTTTKGELKVLMEKELPEGH SQRDKSAYQTIDSAEAAADFAPAVPFAVPGGRSQDAR GY STANGAMATIGPLGSWQWERCLSARDGSERILL LLLGSGQGPQQVGAQQTFEYLKREHSLSKP YQGEAPRCTLRDWELQVHRKHGQGKKN HGDGLAWYTKGWQWERCLSARDGSERILL LLLGSGQGPQQVGAQQTFEYLKREHSLSKP YQGEAPRCTLRDWELQVHRKHGQGKNH HGDGLAWYTKGRQFOFFVGMDKFVOLG VFVDTYPNEKQGEKVFPLSAMNNINGSLSV DHERDGRTELIGCTAIVRILHVDTILVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YVFGTSSITGOLSDNHDVISLKLFVERTPE EEKLHRDVFLPSVDDMKILPENTAPLPFLSGL ALFLLYFFSLVFSVFAVIGIIL YNKW QEGSEK RFY AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRI-PRRGLWILLA HILFMTTACQEANYGALLEELLDCTWHING KUCCPPNAEVDRYGALREELLDCTWHING KUCCPPNAEVDRYGALREELDCTWHING KUCCPPNAEVDRYGARSTANDIAGNI KUCCPPNAEVDRYGAGARSTANDIAGNI KUCCPPNAEVDRYGAGARSTANDIAGNI KUCCPPNAEVDRYGAGARSTANDIAGNI				1			LILFOGGKEAMRRPOIDKKGRAVSWTFSEEN
1025 2375 A 8546 2194 1707 TVSPIKETMASI.KCSTVVCVICLEKPKYRCPA CRYPYCSVVCFRKHKBOCAPETRPVEKKIRS ALPITUTVEPVENDDDDSJADFINSDEEDD VSLQNLKNLGESATLRSI.LINPILRQLMVNI DQGEDKAKLMRAYMQEPLFVEFADCCLGIV VSLQNLKNLGESATLRSI.LINPILRQLMVNI DQGEDKAKLMRAYMQEPLFVEFADCCLGIV VGMELPAVNLKVILLGHWLLTTWGCIVFSGS VGMELPAVNLKVILLGHWLLTTWGCIVFSGS VGMELPAVNLKVILLGHWLLTTWGCIVFSGS LLATTIFLDIVHISIFYPRVSI.LDTGRFGVGMANL SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR GY GY GY GY GY GY GY G			1	İ			VIREFNLNELYORAKKLSKAGDNIPEEOPVAS
1026 2376 A 8547 1078 594 VGMELPAVILXVILLUMINITY VGLIVES		(1	İ	ĺ	ĺ	
1026	1025	2375	Α	8546	2194	1707	TVSFHKTMASLKCSTVVCVICLEKPKYRCPA
ALPTKTYKPYENKODDDÍSADFINSDEEEDS							
VSLQNLKNLGESATLRSLLINPHLRQLMVNIL	1			į		ţ	ALPTKTVKPVENKDDDDSIADFLNSDEEEDR
DOGEDKAKLMRAYMQEPLFVEFADCCLGIV	ł	1	l	1		l	VSLONLKNLGESATLRSLLLNPHLROLMVNL
1026 2376	1					1	
YAWANFTILALGVWAVAQRDSIDAISMFLGG LLATIFLDIYHISPYRVSLIDTORGYGMANL SLLLKPLSCCFVYHMYREGGELLVHTGFLG SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR GY GY GY GY GY GY GY G	1						
	1026	2376	Α	8547	1078	594	
LLATIFLDIVHISISTYPRVSLTDTGREGGGMANL	ļ	1	l				
1027 2377 A 8557 1 340 DFLGPASPQEEGSESSTMTELETAMGMIDV							
SSQDRSAYQTIDSAEAPADPFAVPEGRSQDAR GY							SLLLKPLSCCFVYHMYRERGGELLVHTGFLG
1027 2377 A 8557 1 340 DPLGPASPQEEGGSESSTMTELETAMGMIIDV FSRYSGSEGSTQTLTKGELKVLMEKELPOFLQ SGKDKDAVDKLLLDLDANGDAVDFSEFIVF VAAITSACHKYFEKAGLK 1028 2378 A 8569 20 963 KMAATLGPLGSWQQWRRCLSARDGSRRLLL LLLLGSQQGPQQVGAQQTFEYLKEHGLSKY YQGEAPRCPLRDWELQVIFKIHQGKKNL HGDGLAIWYTKDRMQPGPVFGMDKFVGLG VFPDTYPNEEKQQERVFPYISAMVNNGSLSY DHERDGRTELGGCTAIVRNLHYDTFLVIRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNFDVISLKLFELTVERTPE EEKLHRDVFLPSVPAIVGIILYNKWQEQSRK RFY 1029 2379 A 8572 1 578 AAAASHRSRARSRPRVSSGPAPRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLPMTTACQEANYGALRELCLTQFQVDM EAVGETL WCDWGRTRSYRELADCTWHMAE KLGCFWPNAEVDRFRAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTILLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLITTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPFLLPLLLLQPPAPLQPGGSTGSGP SYLYGVTQFKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRKGHFHGQSFYSTRPPSIHKDY WNRLFILWTEGGESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLITQAVTIT TTWRPSSTTTIAGLRVEKGHSESWHSLDT AIRVALAVAVLKTVILGLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCPPKPGRTRHALCSVVLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLUS EQQQSTEPEQAELOPPUNHYLKLIDRSRALYYKLUS EQQQSTEPEQAELOPPUNHYLKLIDRSRALYYKLUS EQQQSTEPEQAELOPPUNHYLKLIDRSRALYYKLUS EQQQSTEPEQAELOPPUNHYLKLIDRSRALYYKLUS EQQQSTEPEQAELOPPUNHYLKLIDRSRALYYKLUS EQQQSTEPEQAELOPPUNHYLKLIDRSPRVDI VDAKLKIPVSGSKSEGILYVHSKGGGFGORVKE 1032 2382 A 8593 2558 961 RRRFRILPGAEPCEPKVGGRRADMGCSAKAR	1						
FSRYSGSEGSTQTLTKGELKVLMEKELPGFLQ SGKDKDAVDKLLKDLDANGDAQVDFSEFIVF VAAITSACHKYPEKAGIK 1028 2378		<u> </u>	1				
SGKDKDAVDKLLKDLDANGDAQVDFSEFIVF	1027	2377	A	8557	1	340	DFLGPASPOEEGGSESSTMTELETAMGMIIDV
SGKDKDAVDKLLKDLDANGDAQVDFSEFIVF							FSRYSGSEGSTOTLTKGELKVLMEKELPGFLO
VAAITSACHKYPEKAGI.K 1028 2378 A 8569 20 963 KMAATLGPLGSWQWRCLSARDGSRILLL LLLLGSGQGPQQVGAGQTFEYLKREHSLSKP YQGEAPRPCFLRDWELQVHFKHIGQGKKNL HGDGLAWYTKDMOPGPYFGNMDKFVGLG VFVDTYPNEEKQQERVFPYISAMVNNGSLSY DHERDGRPTELGGCTAIVRNIHYDTELVTRY VRRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVPLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFISLVFSVFAIVIGIILYNKWQEQSRK RFY 1029 2379 A 8572 I 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGL WILLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSLTYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLITTOALROPHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQQCGSTGSGP SYLYGVTQFKHLSASMGGSVEPFSFYYPWEL ALVIPNVRISWRRGHROGSYSTRPSHIKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWFRSSTTITAGLRYTESKGHSESWHLSLDT ARVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGTYPCFFKPGRTRHALCSVVLLLLTGGL AFDDFQESCAMMWQKYAGSRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNHYLKLIDRENFYDI VDAKLKIPVSGSKSEGLLYYHISSRGGPFQRW HLDEVFLELKDGQDIPVFKLSGENGDEVKE 1032 2382 A 8593 2558 961 RRTPRLLFGAEPCEPRVGPRRADMGCSAKAR		 -					SGKDKDAVDKLLKDLDANGDAOVDFSEFIVF
LILLIGSOQGPQQVGAGQTTFEYILKREHSLSKP YQGEAPRPCFLRDWELQVHFKHIGQGKKNL HGDGLAIWYTKDRMQPGPVFGNMDKPVGLG VFVDTYPNEEKQQERVFPYISAMYNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVIRY VKRH.TIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAVIGIIL YNK WQEQSRK RFY 1029 2379 A 8572 I 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETILWDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILVPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLQPAPLQPGCCWSPRPLEK NKAMGRPLLLPLLLLQPPAFLQPGSTGSGP SVLYGVTQFKHLASSMGGSVEIPFSFYYPWEL ALVPNVRISWRGHFHGQSFYSTRPPSHKDY VNRLFLNWTEGQESGFIRSINLRKEDGSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT ARVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMVQKYAGSRRSMPLGARIL HIGVFVAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYYHSSRGGFPGRW HLDEVFLEKLGGQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					•		
YQGEAPRPCFLRDWELQVIFFKIHGQGKKNL HGDGLAIWYTKDRMQPGPVFGNMDKFVGLG VFVDTYPNEEKQQERVFPYISAMYNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVRRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY 1029 2379 A 8572 I 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLPMTTACQEANYGALLRELCLTOPQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALLVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSQH SYLYGVTOPKHASASMGGSVEIPPSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPSHKDY VNRLFLNWTEGQESGFLISINLKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLITTQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT ARVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQXYAGSRRSMPLGARIL FHGYFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNHYLKLIDRENFVDI VDAKI.KIPVSGSKSEGLLYVHSSRGGFFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1028	2378	A	8569	20	963	KMAATLGPLGSWQQWRRCLSARDGSRRLLL
YQGEAPRPCFLRDWELQVIFFKIHGQGKKNL HGDGLAIWYTKDRMQPGPVFGNMDKFVGLG VFVDTYPNEEKQQERVFPYISAMYNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVRRY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY 1029 2379 A 8572 I 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLPMTTACQEANYGALLRELCLTOPQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALLVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSQH SYLYGVTOPKHASASMGGSVEIPPSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPSHKDY VNRLFLNWTEGQESGFLISINLKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLITTQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT ARVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQXYAGSRRSMPLGARIL FHGYFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNHYLKLIDRENFVDI VDAKI.KIPVSGSKSEGLLYVHSSRGGFFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR							
HÖDGLAIWYTKDRMOPGPVFGNMDKFVGLG VFVDTYPNEEKQQERVFYJSAMDKFVGLG VFVDTYPNEEKQQERVFYJSAMDKFVGLG VFVDTYPNEEKQGERVFYJSAMDKFVGLGVFVPYSAMDKFVGLG VFVDTYPNEEKQGERVFYJSAMDKFLPGFTVERYP VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVPLPSVDNMKLIPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY 1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGL WLLLA HHLFMITACQEANYGALLRELCLTOPQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGGLCWSPRPLEK NKAMGRPLLPLLLLLQPPALQPGGSTGSQP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL ALVPNVRISWRRGHFHQQSFYSTRPPSIHKDY VNRLFLNWTEGGESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLITIQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT ARVALAVAVLKTVILGLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVULLLTGQL AFDDFQESCAMMWQYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGFPGRW HLDEVFLELKDGQQIPVKKLSGENDDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	•						
VFVDTYPNEEKQQERVFPYISAMVNNGSLSY DHERDGRPTELGGCTAIVRNLHYDTFLVIRTY VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGDLSDNHDVISLKLFELTVERTPE EEKLHRDVFLPSVDNNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY 1029 2379 A 8572 I 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRGL WLLLA HHLPMTTACQEANYGALLRELCLTOPQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRTISLCTSAKSSWTYRSGRLSWQSIKG THILITTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLPLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTLAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGTYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGFPQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGGEPCEPRVGPRRADMGCSAKAR							
VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG YYFGTSSITGOLSDNI-DVISLKLFELTVERTTE EEKLHRUVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAVIGIILYNKWQEQSRK RFY 1029 2379 A 8572 I 578 AAAASHRSRARSRPRRVSSGPAPRAQSSAG RVASGLDSAPLCTMARALCRLPRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSLVPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTGGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRGHFHGQSFYSTRPPSHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLICLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGYFYAGGFAIVYYLQKFHISRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGGSKSGLLLYVHSSRGGPFQRW HLDEVFLELKDGQQDPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR							
1029 2379 A 8572 I 578 AAAASHRSRARSPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRIPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTTRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLITITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNHHYLKLDRENFYDI VDAKLKIPVSGKSEGLLYVHSSRGGPRW HLDEVYFLELKDQQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	i		}				
EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY			ŀ	1			VKRHLTIMMDIDGKHEWRDCIEVPGVRLPRG
EEKLHRDVFLPSVDNMKLPEMTAPLPPLSGL ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK RFY							YYFGTSSITGDLSDNHDVISLKLFELTVERTPE
1029 2379 A 8572 1 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSLLYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLDPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPLNIHYLKLIDRENFYDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE							
1029 2379 A 8572 I 578 AAAASHRSRARSRPRRVSSGPAPRRAQSSAG RVASGLDSAPLCTMARALCRLPRRGLWLLLA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPLNHYLKLIDRENFYDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE				i			ALFLIVFFSLVFSVFAIVIGIILYNKWQEQSRK
RVASGLDSAPLCTMARALCRLPRRGLWILLIA HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETL WCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGYFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGILYVHSSRGGPFQRW HLDEVFLELKDGQQUPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR							
HHLFMTTACQEANYGALLRELCLTQFQVDM EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPQQLCWSPRPLEK NKAMGRPLLLPLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENIFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1029	2379	A	8572	1	578	
EAVGETLWCDWGRTIRSYRELADCTWHMAE KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRSNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR				1			
KLGCFWPNAEVDRFFLAVHGRYFRSCPISGR AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLS WQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR							
AVRDPPGSILYPFIVVPITVTLLVTALVVWQS KRTEGIV 1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTTTQALRQPLHRAPLLPQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	!						EAVGETLWCDWGRTIRSYRELADCTWHMAE
1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTTQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE							
1030 2380 A 8574 1352 372 DSSTVKGGSESRHLCLIPDLKGKARTREASSG SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGFFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE							
SRTCGRRTSLCTSAKSSWTYRSGRLSWQSIKG THLTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1020	0200			10.55	0.550	
THILTITQALRQPLHRAPLLPGQLCWSPRPLEK NKAMGRPLLLPLLILLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1030	2380	A	8574	1352	372	
NKAMGRPLLLPLLLLQPPAFLQPGGSTGSGP SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					l		
SYLYGVTQPKHLSASMGGSVEIPFSFYYPWEL AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFYDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR			-				
AIVPNVRISWRRGHFHGQSFYSTRPPSIHKDY VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					·		
VNRLFLNWTEGQESGFLRISNLRKEDQSVYF CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					l		
CRVELDTRRSGRQQLQSIKGTKLTITQAVTTT TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR		,					
TTWRPSSTTTIAGLRVTESKGHSESWHLSLDT AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR							
AIRVALAVAVLKTVILGLLCLLLLWWRRRKG SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR							
SRAPSSDF 1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					ļ		
1031 2381 A 8580 905 340 RRTAGIYPCFPKPGRTRHALCSVVLLLLTGQL AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					ì		
AFDDFQESCAMMWQKYAGSRRSMPLGARIL FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	100:	0001					
FHGVFYAGGFAIVYYLIQKFHSRALYYKLAV EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR	1031	2381	A	8580	905	340	
EQLQSHPEAQEALGPPLNIHYLKLIDRENFVDI VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					1		
VDAKLKIPVSGSKSEGLLYVHSSRGGPFQRW HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					i	ľ	
HLDEVFLELKDGQQIPVFKLSGENGDEVKKE 1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR				1		ļ	
1032 2382 A 8593 2558 961 RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR					ļ	ł	
idea and other controller in the controller in the controller in the control in t							HLDEVFLELKDGQQIPVFKLSGENGDEVKKE
	1032	2382	Α	8593	2558	961	RRRPRLLPGAEPCEPRVGPRRADMGCSAKAR
							WAAGALGVAGLLCAVLGAVMIVMVPSLIKQ

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
			1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1 ·		1	ł	peptide		/=possible nucleotide deletion, \-possible
			<u> </u>	sequence		nucleotide insertion QVLKNVRIDPSSLSFNMWKEIPIPFYLSVYFFD
			·			VMNPSEILKGEKPQVRERGPYVYREFRHKSNI
					'	TFNNNDTVSFLEYRTFQFQPSKSHGSESDYIV
						MPNILVLGAAVMMENKPMTLKLIMTLAFTTL
						GERAFMNRTVGEIMWGYKDPLVNLINKYFP
						GMFPFKDKFGLFAELNNSDSGLFTGFTGVONI
					,	SRIHL VDK WNGLSK VDF WHSDQCNMINGTS
						GQMWPPFMTPESSLEFYSPEACRSMKLMYKE
{			ł	i		SGVFEGIPTYRFVAPKTLFANGSIYPPNEGFCP
						CLESGIQNVSTCRFSAPLFLSHPHFLNADPVL
						AEAVTGLHPNQEAHSLFLDIHPVTGIPMNCSV
			(KLQLSLYMKSVAGIGQTGKIEPVVLPLLWFA
						ESGAMEGETLHTFYTQLVLMPKVMHYAQYV
						LLALGCVLLLVPVICQIRSQEKCYLFWSSSKK
						GSKDKEAIQAYSESLMTSAPKGSVLQEAKL
1033	2383	Α	8595	595	767	AHLPDTLLLPPHSPTVPTPKSFQCSQKACFSRS
1024	0204		0.505			FCLLLSLVSSSLVSLSLCPPLTQA
1034	2384	Α	8597	640	164	VTTSCIIPFAFGLGVRASERLAEIDMPYLLKYQ
						PMMQTIGQKYCMDPAVIAGVLSRKSPGDKIL
1						VNMGDRTSMVQDPGSQAPTSWISESQVFQTT EVLTTRITELQRRFPTWTPDQYLRGGLCAYSG
						GAGYVRSSQDLSCDFCNDVLARAKYLKRHG
1						F
1035	2385	A	8603	936	204	AMASTLEYSPSPLRRLVGPAAGFSRAARADL
1100						SWDPMAFFTGLWGPFTCVSRVLSHHCFSTTG
İ						SLSAIQKMTRVRVVDNSALGNSPYHRAPRCI
	i			,		HVYKKNGVGKVGDQILLAIKGQKKKALIVG
1						HCMPGPRMTPRFDSNNVVLIEDNGNPVGTRI
				•		KTPIPTSLRKREGEYSKVLAIAQNPV
1036	2386	Α	8606	1	562	PTRAHSFDLCCSPCRRRLLGREEAGEEPTSPV
				!		TQYLQPRSPEECKMFACAKLACTPSLIRAGSR
1		1				VAYRPISASVLSRPEASRTGEGSTVFNGAQNG
						VSQLIQREFQTSAISRDIDTAAKFIGAGAATVG
						VAGSGAGIGTVFGSLIIGYARNPSLKQQLFSY
1037	2207	Α	0616		22/4	AILGFALSEAMGLFCLMVAFLILFAM
103/	2387	Α	8615	2	2364	SPGPSLPESAESLDGSQEDKPRGSCAEPTFTDT
1						GMVAHINNSRLKAKGVGQHDNAQNFGNQSF
						EELRAACLRKGELFEDPLFPAEPSSLGFKDLG PNSKNVQNISWQRPKDIINNPLFIMDGISPTDI
						CQGILGDCWLLAAIGSLTTCPKLLYRVVPRG
						QSFKKNYAGIFHFQIWQFGQWVNVVVDDRL
						PTKNDKLVFVHSTERSEFWSALLEKAYAKLS
] ;						GSYEALSGGSTMEGLEDFTGGVAQSFQLQRP
]		-		i	-	PONLLRLLRKAVERSSLMGCSIEVTSDSELES
]						MTDKMLVRGHAYSVTGLQDVHYRGKMETLI
j						RVRNPWGRIEWNGAWSDSAREWEEVASDIQ
						MQLLHKTEDGEFWMSYQDFLNNFTLLEICNL
]					.	TPDTLSGDYKSYWHTTFYEGSWRTGSSAGGC
1						RNHPGTFWTNPQFKISLPEGDDPEDDAEGNV
ļ i						VVCTCLVALMQKNWRHARQQGAQLQTIGFV
						LYAVPKEFQNIQDVHLKKEFFTKYQDHGFSEI
[FTNSREVSSQLRLPPGEYIIIPSTFEPHRDADFL
						LRVFTEKHSESWELDEVNYAEQLQEEKVSED
						DMDQDFLHLFKIVAGEGKEIGVYELQRLLNR
						MAIKFKSFKTKGFGLDACRCMINLMDKDGSG
						KLGLLEFKILWKKLKKWMDIFRECDQDHSGT
						LNSYEMRLVIEKAGIKLNNKVMQVLVARYA DDDLIIDFDSFISCFLRLKTMFTFFLTMDPKNT
						GHICLSLEOVLGEGWEGICRIAPACPSTPPPPS
L	ــــــا	L	ــــــا			GUICESPECA FORCAM VCI STILLIS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	l	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	{		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1			amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
Ì	l	1		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
		l		peptide	· -	/=possible nucleotide deletion, \=possible
				sequence	ł	nucleotide insertion
						SDVPGPASCPRLFPPWDLLPVSTVAADDHVGI
1						EAL
1038	2388	Α	8621	3	1494	RSRMARAPLGVLLLLGLLGRGVGKNEELRLY
		1				HHLFNNYDPGSRPVREPEDTVTISLKVTLTNL
		1	1			ISLNEKEETLTTSVWIGIDWQDYRLNYSKDDF
	-					GGIETLRVPSELVWLPEIVLENNIDGQFGVAY
	l	ļ				DANVLVYEGGSVTWLPPAIYRSVCAVEVTYF
1						PFDWQNCSLIFRSQTYNAEEVEFTFAVDNDG
1		ŀ				KTINKIDIDTEAYTENGEWAIDFCPGVIRRHH
	İ	ì				GGATDGPGETDVIYSLIIRRKPLFYVINIIVPCV
Ì		1				LISGLVLLAYFLPAQAGGQKCTVSINVLLAQT
						VFLFLIAQKIPETSLSVPLLGRFLIFVMVVATLI
		ł				VMNCVIVLNVSQRTPTTHAMSPRLRHVLLEL
ł	l	1	ł	}		LPRLLGSPPPPEAPRAASPPRRASSVGLLLRAE
			1			ELILKKPRSELVFEGQRHRQGTWTAAFCQSL
İ			1			GAAAPEVRCCVDAVNFVAESTRDOEATGEE
1						VSDWVRMGNALDNICFWAALVLFSVGSSLIF
						LGAYFNRVPDLPYAPCIOP
1039	2389	A	8636	1	900	PGRERPGGGGARRRPQHLPALLPSERPDCATL
1.00		' '	0000	-] ***	QAMENELPVPHTSSSACATSSTSGASSSSGCN
						NSSSGSGRPTGPQISVYSGIPDRQTVQVIQQ
i		1				ALHRQPSTAAQYLQQMYAAQQQHLMLQTA
1		ì				ALQQOHLSSAQLQSLAAVQQASLVSNRQGST
						SGSNVSAQAPAQSSSINLAASPAAAQLLNRA
[ĺ		•		QSVNSAAASGIAQQAVLLGNTSSPALTASQA
						QMYLRAQMLIFTPTATVATVQPELGTGSPAR
i	1	}		•		PPTPAQVQNLTLRTQQTPAAAASGPTPTQPVL
1]					PSLALKPTPGGSQPLPTPA
1040	2390	Ā	8645	98	1388	ASQLAFGGKLTSTPSRDFQGCGRGAVTCCSF
		_				HEHRHOSGRCLSTGMAPNLKGRPRKKKPCPO
İ			1			RRDSFSGVKDSNNNSDGKAVAKVKCEARSA
						LTKPKNNHNCKKVSNEEKPKVAIGEECRADE
1						QAFLVALYKYMKERKTPIERIPYLGFKQINLW
					ĺ	TMFQAAQKLGGYETITARRQWKHIYDELGG
						NPGSTSAATCTRRHYERLILPYERFIKGEEDKP
						LPPIKPRKQENSSQENENKTKVSGTKRIKHEIP
		}]]	KSKKEKENAPKPQDAAEVSSEQEKEQETLISQ
	1					KSIPEPLPAADMKKKIEGYQEFSAKPLASRVD
1	[PEKDNETDQGSNSEKVAEEAGEKGPTPPLPSA
1	1					PLAPEKDSALVPGASKQPLTSPSALVDSKQES
Ì						KLCCFTESPESEPQEASFPRLPHHTGHRWQTR
L	<u> </u>	_ :				MRRRMTNCPPWQITLPTAP
1041	2391	A	8646	113	1492	LLQEMCTKTIPVLWGCFLLWNLYVSSSQTTYP
			" ' -			GIKARITQRALDYGVQAGMKMIEQMLKEKK
				1		LPDLSGSESLEFLKVDYVNYNFSNIKISAFSFP
1	1					NTSLAFVPGVGIKALTNHGTANISTDWGFESP
ľ						LFVLYNSFAEPMEKPILKNLNEMLCPIIASEVK
	-					ALNANLSTLEVLTKIDNYTLLDYSLISSPEITE
						NYLDLNLKGVFYPLENLTDPPFSPVPFVLPER
			· .			SNSMLYIGIAEYFFKSASFAHFTAGVFNVTLS
	[TEEISNHFVQNSQGLGNVLSRIAEIYILSQPFM
					,	VRIMATEPPIINLQPGNFTLDIPASIMMLTQPK
]					NSTVETIVSMDFVASTSVGLVILGORLVCSLS
						LNRFRLALPESNRSNIEVLRFENILSSILHFGVL
						PLANAKLQQGFPLPNPHKFLFVNSDIEVLEGF
						LLISTDLKYETSSKQQPSFHVWEGLNLISRQW
					İ	RGKSAP
1042	2392	A	8672	538	170	ARRIARTRESKAAVSQDNVPALQPGKKKKLR
l	ı -		l -			LGGKKKKFKFFRLPKEFKKQLMYSPSNFKKM
l						

NO. of No. of India and India of India	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence cotide seq uence						1	
Sequence	nucl-	pentide					
Sequence	1						
uence	1						
amino acid residue of peptide sequence sequence period sequence p	1 .	1	1				
Pesidue of poptide Sequence Y=Tyronine, X=Unknown, *=Siop coden, P-possible unclotide delicito, *p-possible unclotide distriction TSLAGNYTOCLNKLKYVIYSAQYPAYGNITT LDMTSTIDHVLEQDFWICTFYSWKERQ1 TSLAGNYTOCLNKLKYVIYSAQYPAYGNITT LDMTSTIDHVLEQDFWICTFYSWKERQ1 TSLAGNYTOVAL		l] * * * * * * * * * * * * * * * * * * *			
pepide		1					
1043 2393 A 8688 359 17	}					00400000	
1043 2393 A 8688 359 17	j		ļ			ļ	
1043 2393 A 8688 359 17							
1043 2393 A 8688 359 17	1	1		1			
GIMTSLIKPIKRRWRDVKRWXSGGFTGESG HADTILGRGGLOGDISSLLQWOKRILRTE GEPSFKYISKNIPPICSYITGFL 1044 2394 A 8718 292 1490 GIVKTSVATIGHGGSCGOVIQVKSPATQIS GFKFTSKMEDPYMESDSSGOVIQVKSPATQIS GFKFTSKMEDPYMESDSSGOVIQVKSPATQIS GFKFTSKMEDPYMESDSSFDDFWKGEDLSNYS YSSTLPPILLDAAPCEPESLEINKYFVUITYAL VPLISLLGNSLVMLVILSVSVGRSYDDVYLL NIAALADLLFALTLPIWAASKVNGWIFGITELC KVVSLLKEVNFYSGILLLACISVORYLAIVHA TRITLTOKRYLVKFICLSIWGISLLLALPVILER RTVYSSVSPACYEDMONITAWMRMLRIL POGRGFIVPLLIMLFCVGFTLRTLFKAHMGOK HAMRVIFAVVILELUCWPYNILVLADTLM RTOVIQETCERRNHIDRALDATEILGILHSCLN PLIYAFIOQEFRIGLIKILAHHGLISKDSLPKDS RSPVOSSGGITSTIL PROSSGGITSTIL PROSSGGITSTIL PROSSGGITSTIL HIGSNVIQDIETGAFHGLRGLRRLHLINNINKL ELRDDTFIGLENLEYLQVDYNYISVEPNAF GKIHLLQVILINDNLLSSLPNLFFYPUTHL DIRGNKLLIPVGLOHMKVVELQLEEN PWNCSCELISLKDWLDBISYSALVGDVVCETF FRLHGRDLDEVSKOELCPRRLISDYEMRPQTP LISTTOYLHTTTASVNSVATSSSAVYKPPLKPP KGTRQPNKRPVRPTSRQPSKOLGYSNYOPSIA VQTKSPVPLECPTACSCOLGISLIGLINNINOQE RKIESLAELQPKFYNPFKKMYLTENVIAVVRGT DLEATGLDLHLHGNNINSMIQDRAFGDLTN LRRLYLNGNRIERLSPLFYGLQSLQVILFLQY VLREGGGTFDYPNLQLHENNINLOQAMPS GVFSGLTLIRLINLRSNHFTSLPVSGVLDQLKS LIQDLHDNPWDCTCOLGYMKLWVGQLKG VLVDEVICKAPKKPAETDMRSIKSELLCPDYS VNTERFOLDSHLTGMARNSMIGDRAFGDLTN LRRLYLNGNRIERLSPLFYGLQSLQVILFLQY VLVDEVICKAPKKPAETDMRSIKSELLCPDYS VNTERFEDLLSPYQDARRSVALENSTGA PASLGAGGASSVPLSVILLISLLIVFRMSVENA AGLEFULMMRRKKNOSDHTSTNNSDVSSFN MQYSVYGGGGTTGGRPHAHVHIRGPALPK VKTRAGHVYEYPIPPLGBMCCNPITKSRGEN SVEDYYKDLEILKVTYSSNHLDQCQOPPPP QQPQQOPPPOLQLQPGEERRESHHLRSPAYS VSTERFEDLLSPYGDARRSVALENSTGA PASLGAGGASVPLSVYLISSILLIVFRMSVEYN VSTERFEDLLSPYGDARRSVALERSTRINSTDARPH QVLHGAGDSRLREPVLYSPRANEYPRNE YLELKAKLNYEPDVLEVLEKQTTFSOF PSPAAGGLAWVSALLGSGSGRODTSCRGVT AMAGALVRKAADVYRSKDFROYLMSTHFW GPVANWGIPLANDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACARATNEVA QUIGGGRIGHEMTKTASA 1047 2397 A 8741 673 924 ALPOTTPCQTVILLTIORISGRGMOTSSTSNNIPLPP AKFFTSLOSLNWSSHLPFSQVCKRONAK PFTTKLLISSPLWNFFAQQL	1043	2393	A	8688	359	17	
HHADTLGDRGGLQGDHSELLQWQKRLIRTE GEPSFKYISKINPICSYTIGE GEPSFKYISKINPICSYTIGE GEPSFKYISKINPICSYTIGE GEPSFKYISKINPICSYTIGE GEPSFKYISKINPICSYTIGE GEPSFKYISKINPICSYTIGE GEPSFKYISKINPICSYTIGE GEPSFKYISKINPICSYTIGE GEPSFKYISKINPICSYTIGE GEPSFKYISKINPICSYTIGE GEPSFKYISKINPICSYTIGE GEPSFKYISKINPICSYTIGE GEPSFKYISKINPICSYTIGE GEPSFTKYISKINPICSYTIGE GEPSFTKYISKINPICSYTIGE GEPSFTKYISKINPICSYTIGE GEPSFTKYISKINPICSYTIGE GEPSFTKYISKINPICSYTIGE CHARLEY CANADILISTIC CHARLEY CANADIL	1		l ·-		1 337	1 *′	
GEPSKYISKNIPPICSYITIGEL	1		ļ			1	
1044 2394 A 8718 292 1490 GTVKTSVAIPTTAGESCSSGOVQVKSSPATOS GEKFTSKMEDENMESDSEEDEN GELSVYS GTKFTSKMEDENMESDSEEDEN GELSVYS YSSTLPPFLLDAAPCEPESLEINKYFVVIIYAL VFLLSLLONSLYMLVILYSRVORSVTDVYLL NA LAADLLFALTLPIWAASKVNOWIFOTFLC KVVSLLKEVNFYSGILLLAGISVDRYLAIVHA TRITLTQKRYLVKFICLSIWGISLLAPULLFR RTVYSSNVSPACYEDMGNNTANWRMLELD PQSFGFIVPLLDMIFCYGFTLRTIKAHMGOK HRAMRVIPALVIFLICWLPYNLVLLADTLM RTOVJGETCERRINDBALDATEIGLHSCLN PLIVAFIGGOKFHGLLKILAHGLISKDSLFKDS RPSFVSSSGITSTITL GLINECUL PURAFIGAGOKMHTCCPPVTLEQ DLHRKMHSWMLQTLAFAVTSLVLSCAEDISSES PPRFFIYHLLSGNLLNRLYPNGFVYTTGASTLLSCAEDISSES PPRFFIYHLLSGNLLNRLYPNGFVYTTGASTLLSCAEDISSES PPRFFIYHLLSGNLLNRLYPNGFVYTTGASTLLSCAEDISSES PPRFFIYHLLSGNLLNRLYPNGFVYTTGASTLLSCAEDISSES PPRFFIYHLLSGNLLNRLYPNGFVYTTGASTLLSCAEDISSES PPRFFIYHLLSGNLLNRLYPNGFVYTTGASTLLSCAEDISSES PPRFFIYHLLSGNLLNRLYPNGFVYTTGASTLLSCAEDISSES PPRFFIYHLLSGNLLNRLYPNGFVYTTGASTLLSCAEDISSES PPRFFIYHLLSGNLLNRLYPNGFVYTTGASTLLSCAEDISSES PPRFFIYHLLSGNLLNRLYPNGFVYTTGASTLLSCAEDISSES PPRFFIYHLLSGNLLNRLYPNGFVYTTGASTLLSCAEDISSES PPRFFIYHLLSGNLLNRLYPNGFVYTTGASTLLSCAEDISSES PPRFFIYHLSGNLLSCAEDISSES PPRFFIYHLSGNLLSCAEDISSES PPRFFIYHLSGNLLSCAEDISSES PPRFFIYHLSGNLLSCAEDISSES PPRFFIYHLSGNLLSCAEDISSES PPRFFIYHLSGNLLSCAEDISSES PPRFFIYHLSGNLLSCAEDISSES PPRFFIYHLSGNLLSCAEDISSES PPRFFIYHLLSGNLLSCAEDISSES PPRFFIYHLSGNLLSCAEDISSES PPRFFIYHLSGNLSCAEDISSES PPRFFIYHLSGNLSCAEDISSES PPRFFIYHLSGNLSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSGNLTSCAEDISSES PPRFFIYHLSCAEDISSES PPRFFIYHLSCAEDISSES PPRFFIYHLSCAEDISSES PPRFFIYHLSCAEDISSES PPRFFIYHLSCAEDISSES PPRFFIYHLSCAEDISSES PPRFFIYHLSCAEDISSES PPRFFIYHLSCAEDISSES PPRFFIYHLSCAEDISSES PPRFFIYHLSCAEDISSES PPR]			
GFKFTSKMEDFNMESDSEEDFWXGEDL SNYS YSSTLPPFLLDAAPCEPESLEINKYFVVIIV SNYS YSSTLPPFLLDAAPCEPESLEINKYFVVIIV SNYS YSSTLPPFLLDAAPCEPESLEINKYFVVIIV SNYS VYELISLIGNSLYMLIVILYSRVORSYTDYYLL NIAALADLIFALTI-PIWAASKYNOWIFGTEL KVYSLKEVINYSSGLLACISVITYHA TRITLTQKRYL-VKFICLSIWGLSLLALPVLITHA TRITLTQKRYL-VKFICLSIWGLSLLALPVLITHA RTQVIQETCERRNHIDRALDATELIGILHSCLIN PLYAFIGQKFRHGLLKILAHGLISKDSLFKDS RSFVOSSSGHTSTTIL PLYAFIGQKFRHGLLKILAHGLISKDSLFKDS RSFVOSSSGHTSTTIL DLHAFNISWALDATELIGILHSCLIN PLYAFIGGKRHGLKILAHGLISKDSLFKDS RSFVOSSSGHTSTTIL HLGSNVIQDIETGAFHGLKRILAHGLISKDSLFKDS PRPFIPHLLISGNUTALFAYTSLVLSCAETIDY YGEICDNACPCEEKDGILTYSCENRGISISSEIS PRFFIPHLILISGNUTALFAYTSLVLSCAETIDY YGEICDNACPCEEKDGILTYSCENRGISISSEIS PRFFIPHLLISGNUTALFAYTSLVLSCAETIDY YGEICDNACPCEEKDGILTYSCENRGISISSEIS PRFFIPHLILISGNUTALFAYTSLVSCAETIDY YGEICDNACPCEEKDGILTYSCENRGISISSEIS PRFFIPHLILISGNUTALFAYTSLVSCAETIDY YGEICDNACPCEEKDGILTYSCENRGISISSEIS PRFFIPHLILISGNUTALFAYTSCHAFT GILHATGLIDLRYLQVDYNTIVGASI HLGSNVIQDIETGAFHGIRGLRRILHINNNKL ELLRDDTFLOLENYLQVDYNTIVGASI PWKCSCELISLKDWLDSISYSALVODVVCETP FRIHGRDLDEVSKQELCPRRILISDYEMRPOTP FRIHGRDLDEVSKQELCPRRILISDYEMRPOTP LSTTGYLHTTPASVINSVASTSSSAVTOVVCETP FRIHGRDLDEVSKQELCPRRILISDYEMRPOTP FRIHGRDLDEVSKQELCPRRILISDYEMRPOTP LSTTGYLHTTPASVINSVASTSSAVTOVVCETP FRIHGRDLDEVSKQELCPRRILISDYEMRPOTP LSTTGYLHTTPASSIPPLYGLQCIALFAYNACPRATIALVAVRRT DLLEATGJLLIHLIGNNRISMIQDRAFGDLIT LRRLYVLONNRERSPEPLYGLGSLQTIFLOV NILBERGSGTTDPVPNLQLLIFLNNNLLQAMPS GVFSGLTILIRILKRSNHFTSLPDEVJQLQKS LIQDIHDHPWDCTCDIVGMKLWVEQLKV VVDEPCIKAPKKATGAETHANVALVARA PASLGAGGGASSVFLVILSILLVFIMSVEVA AGLFFVLVMKRRKKNQBHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHHRGPALDF VXTPAGHYYEVIPHCIGHCKNSTNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALDF VXTPAGHYYEVIPHCIGHCKNSTNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALDF VXTPAGHYYEVIPHCIGHCKNSTNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALDF VXTPAGHYYEVIPHCIGHCKNSTNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALDF VXTPAGHYYEVIPHCIGHCKNSTNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALDF VXTPAGHYYEVIPHCIGHCKNSTNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALDF VXTPAGHYYEVIPHCIGHCKNSTNSDVSSFN MQYSVYGGGGGTGGHPHAHVHR	1044	2394	A	8718	292	1490	
YSSTLPFFILDAAPCEPESLEKKYEVUIYAL VFLISLIGINSLVAN, VILYSRVGRSVTDVYLL NLALADLIFALTPIWAASKVNGWIFGTFLC KVVSLIKEVNFYSGILLACISVDRYLAUPULFR RTITQKRYLVKFICISWIGISLIAPULFILFR RTIYGSSNVSPACYEDMGNNTANWRMLILEIL POSFOFIVPILIMIFCYGFITERTIFKAHMGOK HRAMRVIFAVVLIFLLCWIPYNLVLLADTIL RTOYUGETCERRNIMDRALDATELIGILHSCLN PLIYAFIGGKFRHGILKILAHGLISKDSLPKDS RPSFYGSSSGHTSTTI TOYUGFICERNIMDRALDATELIGILHSCLN PLIYAFIGGKFRHGILKILAHGLISKDSLPKDS RPSFYGSSSGHTSTTI TOYUGFICERNIMDRALDATELIGILHSCLN PLIYAFIGGKFRHGILKILAHGLISKDSLPKDS RPSFYGSSSGHTSTTI TOYUGFICERNIMPALORISSISSIS PRFRIYHLLSGNLINRLYPNEFVYTTGASL HLGSNVQDIETGAFHGLGILRELHINNIKL ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF GKHLLQVILLNDNLISSIPNALVQUVVCETP FRLIGGDLOVECCEKDGLIVSCENHEYLNINLINGL ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF GKHLLQVILLNDNLISSIPNALVQUVVCETP FRLIGGDLOVECKGELCPRRLISDYEMRQTP LSTTGYHLHTTPASVSVATSSSAVYKEPLEVF KGTRGPNKPRVRTSRGPSCLDLGYSNYGPSIA YQTKSPPPLECTTACSCNLQISDLGLNNVCOC RKESISALQPKPLYNFKKMYLTERYIAVVRRT DLLEATGLDLIHLGNNRISMIGDRAFGGLTI LRALVNGNRIERS,PELYGIG,SLQVLFLQV NLREIGGSTDPYPHCLQLLENNLLQAMPS GVPSGLTLIRLINS.PELYGIG,SLQVLFLQV NLREIGGSTDPYPHCLQLICHNNLLQAMPS GVPSGLTLIRLINS.PELYGIG,SLQVLFLQV NLREIGGSTDPYPHCLQLICHNNLLQAMPS GVPSGLTLIRLINS.PELYGIG,SLQVLFLQV NLREIGGSTDPYPHCLQLICHNNLLQAMPS GVPSGLTLIRLINS.PELYGIG,SLQVLFLQV NLREIGGSTDPYPHCLQLICHNNLLQAMPS GVPSGLTLIRLINS.PELYGIG,SLQVLFLQV NLREIGGSTDPYPHCLQLICHNNLLQAMPS GVPSGLTLIRLINS.PELYGIG,SLQVLFLQV NLREIGGSTDPYPHCLQLICHNNLLQAMPS GVPSGLTLIRLINS.PELYGIG,SCNGTONSOSSN MQVSVYGGGGGTGGHPHAHVHIRGPALP VKTPAGHYYEVPRATALINSTON VKTPAGHYYEVPLCHERQTTYSQF QQQQOPPOLQLOPGEERERSSHLIR.SPAPS QVSVYGGGGGTGGHPHAHVHIRGPALP QVLMPGAGBDSRREPVILSPSREPHILS.PSPAPS QVGQOPPOLQLOPGEERERSSHLIRSPAPS QVGQOPPOLQLOPGEERERSPHILSSPAPS QVSVYGGGGGTGGHPHAHVHIRGPALP QVLMPGAGBDSRREPVILSPSSEN SYSTPENTNERP YLELKAKLNVEPDYLEVLERQTTTSQF GPVANGLPIAANDMKSSFERISISGRMTFALC CVSLTFMRFAYEVVPPNALLFACATINEVA	1000		J	0,10	272	1450	
VFILSILGNSLVMLVIL YSRVGRSTYDYLL NLALADLERALTEIWAASKYNOWFGTFIC KVVSLLKEVNFYSGILLACISVDRYLAIVHA TRITIQKRYLVKFICISWGLSLLLALPVLIFR RIVYSSNVSPACTEDIMGNINAWMALIRIL PQSFGFIVPLIMLFCVGFTLRTLFKAHMGQK HRAMRVIFAVUIFLICUWPLVILLADTIM RTQVIGETCERRINIDRALDATEILGILHSCLN PLIYAFFIGKFRGILKILAIHGLISKDSLPKDS RPSFVGSSSGHTSTTI 1045 2395 A 8724 254 3184 FRANLAITVANRGGAGGKMHTCCPPVTLEQ DLIRKMHISWMQTLARAVTSLVLSCAETIDY YGEICDNACPCEEKDGILTVSCENRGIISLSEIS PREFPIYHLLISGNLINRLIPPNYTGASSL HLGSNVIQDIETGAFFIGLRGLRELHINNINKL ELLRDTFIGLEINLEYLQVDYINYUSEPNAF GKLHLLQVLILNDNLLSLPANLFRFVPLTHL DLRGNRIKLLPYVGLLGHMDKVVELQLEEN PWNCSCELISLKDWLDSISYSALVGDWVCETP FRLHGRDLDEVSKQELCPRRLISDYERRYPTT STTGYJHITTPASVNSVATSSALVGDWVCETP FRLHGRDLDEVSKQELCPRRLISDYERRYPTT LSTTGYJHITTPASVNSVATSSALVGDWVCETP FRLHGRDLDEVSKQELCPRRLISDYERRYPTT LSTTGYJHITTPASVNSVATSSALVGDWVCETP FRLHGRDLDEVSKQELCPRRLISDYERRYPTT LSTTGYJHITTPASVNSVATSSALVGDWVCETP FRLHGRDLDEVSKQELCPRRLISDYERRYPTT LSTTGYJHITTPASVNSVATSSALVGDWVCETP FRLHGRDLDEVSKQELCPRRLISDYERRYPTT LSTTGYJHITTPASVNSVATSSALVGDWVCETP FRLHGRDLDEVSKQELCPRRLISDYERRYPTT LSTTGYJHITTPASVNSVATSSALVGDWVCETP FRLHGRDLDEVSKQELCPRRLISDYERRYPTIF LSTTGYJHITTPASVNSVATSSALVGDWVCETP FRLHGRDLDEVSKQELCPRRLISDYERRYPTIF LSTTGYJHITTPASVNSVATSSALVGBWVCETP FRLHGRDLDEVSKQELCPRRLISDYERRYPTIF LSTTGYJHITTPASVNSVATSSALVGBWVCETP FRLHGRDLDEVSKQELCPRRLISDYERRYPTIF LSTTGYJHITTPASVNSVATSSALVGBWVCETP FRLHGRDLDEVSKQELCPRRLISDYERRYPTI LSTRYPLYBYGJCLQAMPS GVPSSPLITERSVSALVGBROTSGALVGT ANGRALVRKAKRAETDMRSIKSELLCPDYS UVVSTPTPSSIQVPARTSAVTPAVRLINSTGA PASLGAGGGASVPLSVILLSLLLVFINSVFVA AGLFVLWRRRKKNQSDFTSTNNSDVSFN MQYSVYGGGGTGGHHAHVHHRGFALPK VKTPAGHLVEVPHPLGHMCKNPITYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPP QQQQPPPQLQPPPQLQQPPGPLQLQPGEERRSHHLRSPAYS VSTEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLESTYRFPCSPAAVTFSNDLRFHPY QVLHPGAGDSRLREPVLYSPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFNSP SYSTEMPEDLLSPVQDADRFYRGILEPDKHCST TPAGNSLESTYRFPCSPAAVTFSNDLRFHPY QVLHPGAGDSRLREPVLYSPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFNSP GPANNGCIPLAANDMKKSPESIGGRMTFALC CYSLTFMRPAVKVQPRWULFACHATNEVA QLIQGGRLIKHEMTATAS. PTTTKLLH							L
NLALADLEATTLEIWAASKVNGWIFGTFLC KVVSLIKEVRYSGILLACSVPLAIVHA TRTLTQKRYLVKFICLSIWGLSLLALPVLIFR RTYYSSNSPACYEDMGNNTANWRAILRIL PQSFGRIPVLIMLECVGFTLEIGLHSCLN RTQVIQETCERRHIBRALDATIM RTQVIQETCERRHIBRALDATIM RTQVIQETCERRHIBRALDATIM RTQVIQETCERRHIBRALDATIM RTQVIQETCERRHIBRALDATIM RTQVIQETCERRHIBRALDATIM RTQVIQETCERRHIBRALDATIM RTQVIQETCERRHIBRALDATIM RTQVIQETCERRHIBRALDATIM RTQVIQETCERRHIBRALDATIM RTQVIQETCERRHIBRALDATIM RTQVIQETCERRHIBRALDATIM RTQVIQETCHAPTISVLYSCAETIDY YGEICDNACPCEREDGILTVSCRETIBLY YGEICDNACPCEREDGILTVSCRETIBLY STORTHLILBGNILINBLYPNEFVNYTGASIL HLGSNYQDIETGAHFIGIREHLINNNILL ELRDDTFLGLENLEYLQVDYNYISVIEPNAF GKHLLQVILINDNILSSLPNTLFREVPLTHL DIRGNRIKLILPVGLIQHMOKVVELQLEEN PWNCSCELISLKDWLDSISYSALVGDVVCETP FRIHGRDLDEVRYQELYENTERPOTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVRPTSROPSKOLLGYSNYQPSIA YQTKSPVPLECPTACSCNLQISDLGINNXCQE RKIESIAELQPKPYPPKKMYLTENYIAVVRRT DLLEATGLDLLHLGNNINSMQDRAFGDLTN LRRLYLNGRRIERLSPELFYGLOSLQYLFLQY NLIRRISQSGTFDYPNOLQLIFLINNILLQAMPS GVFSG.TLLRLINLRSNHFTSLPVSGVLDGLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKPATDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLINSTGA PASLGAGGGASVPLSVILISLLIVFINSVEVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQVSVYGGGGTGGHPHAHRGPALPK VKTPAGNIFEVRFCSPAATYHRGPALPK VKTPAGNIFEVRFCSPAATYHRGPALPK VKTPAGNIFEVRFCSPAATYHRGPALPK VKTPAGNIFEVRFCSPAATYHRGPALPK VKTPAGNIFEVRFCSPAATYHRGPALPK VKTPAGNIFEVRFCSPAATYHRGPALPK VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLEPVFRFCSPAATYHRGPALPK VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLEPVFRFCSPAATYHRGPALPK VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLEPVFRFCSPAATYFRFCSPAATYHRNE VLELKAKLNVEPDYLVLSEQTFTSQF SPSAAGGLAWVSLALGSGSRGHNSGSGVGT AMAGALVRKAADVVRSKDFRDYLLSTHFW GPVANWGLIPAANDMKKSPESIGSGRHNSGSGVGT AMAGALVRKAADVYRSKDFRDYLMSTHFW GPVANWGLIPAANDMKKSPESIGSGRMTSALC CYSLTFMRPAVKVQPRWILFACHATNEVA QULQGGRILKHEMTKTASA PTTTKLLHSSPYLWFFAQAL CYSLTFMRPATKSQSLNWSSHLIPSSTATESVGKRGNAK PTTTKLLHSSPYLWFFAQAL PARTSCLUBERTYTKLASPATESVGKRGNAK PTTTKLLHSSPYLWFFAQAL PARTSCLUBERTYTKLASPPATESVGKRGNAK PTTTKLLHSSPYLWFFAQAL							I
KVVSLIKEVNFYSGILLACISVORYLATVHA TRTITOKRYJ VKPICLSIWGLSLLALPVLIFR RTYYSSNVSPAC YEDMONNTANWRMLIRIL PQSFGIVPLIMLFCYGFILRILFKAHMGQK HRAMRVIFAVUIFLIC WIPYNIVLLADTIM RTQVIGETCERRINIIDRALDATELIGI.HISCIN PLIYAFIGQKFRGILKILAHHGLISKDSI PKDS RPSFVGSSSGHTSTTI. 1045 2395 A 8724 254 3184 FRANLAITVANRGGAGGKMHTCCPPVTLEQ DIFRKMESWMLQTLAFAVTSJV.SCAETIDY YGEICDNACPCEEKDGILTVSCENRGIISI.SEIS PPRFPIYHLLISGNLIJRILYPRYTTGASIL HLGSNVIQDIETGAFHGLRGILRILHINNNKL ELLRDUTFLGLENLEYLQVDYNYISVIEPNAF GKLHLQVLILINDNLISSEIPNYTTGASIL HLGSNVIQDIETGAFHGLRGILRILHINNNKL ELLRDUTFLGLENLEYLQVDYNYISVIEPNAF GKLHLQVLILINDNLISSEIPNYTGTASIL HLGSNVIQDIETGAFHGLRGILRILHINNNKL ELLRDUTFLGLENLEYLQVDYNYISVIEPNAF GKHLQVLILINDNLISSEIPNYTGTASIL HLGSNVIQDIETGAFHGLRGILRILHINNNKL ELLRDUTFLGLENLEYLQVDYNYISVIEPNAF GKHLQVLINDNLISSEINVYTPSTAL GKHLQVLINDNLISSEINVYTPSTAL YQTKSSVPLECPTACSCNLQISDLGINNYQPSIA YQTKSSVPLECPTACSCNLQISDLGINVXQE RKIESIAELQPKFYNPKKMYLTENYIAVVRGT RKIESIAELQPKFYNPKKMYLTENYIAVVRGT RKIESIAELQPKFYNPKKMYLTENYIAVVRGT RKIESIAELQPKFYNPKKMYLTENYIAVVRQE RKIESIAELQPKFYNPKKMYLTENYIAVVRQE RKIESIAELQPKFYNPKKMYLTENYIAVVRQE RKIESIAELQPKFYNPKKMYLTENYIAVVRQE RKIESIAELQPKFYNPKKMYLTENYIAVVRQE RKIESIAELQPKFYNPKKMYLTENYIAVVRQE RKIESIAELQPKFYNPKKMYLTENYIAVVRQE RKIESIAELQPKFYNPKKMYLTENYIAVVRQE RKIESIAELQPKFYNPKKMYLTENYIAVVRQE RKIESIAELQPKFYNPKKMYLTENYIAVVRQ RVSCGILLIRLINLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCCDYS DVVVSTPTPSSIQVARTSATVARLNSTGA PASLGAGGAGSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNSODITSTNNSUSSIN MQVSVYGGGGTGGHPHAHVHRGPALPK VKTPAGHVYEYIPPICJAMCKNPIYRSREGN SVEDYKDLHELKVYTSSNHHLQQQQPPPPP QQPQQPPPQLQLQPQQDQAPPYGLLQPKHCST TAGMSLEPTYRFPCSPAATTSHVLRSTH VKTPAGSILEPTKPFCSPAATTSHVDLRSTH VKLHAALNUSPPALFERPYDLERTH QYLHPGAGDSRLREPVLYSPPSAVFEDNENE YLELKAKLNVEPPVLSPELSCMTTALC CYSLTFMRPAVKVQPRWULFACHATNEVA QLIQGGRILKHEMTKTASA PETTKLLHSSPYLWFFAQAL CYSLTFMRPAVKVQPRWULFACHATNEVA QLIQGGRILKHEMTKTASA PETTKLLHSSPYLWNFFAQAL PETTKLLHSSPYLWNFFAQAL PTTKLLHSSPLWNFFAQAL PTTKLLHSSPLWNFFAQAL PTTKLLHSPLWFATATSVGKRGNAK PTTTKLLHSSPLWNFFAQAL PTTKLLHSPLWNFF			1	1			
TRILTOKRYLVKFICLSIWGLSLLLALPYLIFR RTVYSSNVSPACYEDMGNNTANWRMLERI PQSFGFIVPLIMLECYGFTLRTLFKAHMGQK HRAMRVIFAVVLIFLLCWLPYNLVLLADTIM RTQVQETCERRHIBRALDATEILGILHSCLN PLIYAFIGOKFRIGLIKILAHGLISKDSLFKDS RPSFVGSSGGTSTITL 1045 2395 A 8724 254 3184 FRANLATIVANRRGAQGGKMHTCCPPVTLEQ DLIRKMHSWMLQTLAFAVTSLVLSCAETIDY YGEICDNACPCEREGIGLTSCENGGISLSEIS PRFFIYHLLLSGNLLNIKLYPNEFVNYTIGASL HLGSNYIQDIETGAFHGURGRHLIMNIKL ELIRDDTFLGLENLEVLQVDYNYISVIEPNAF GKLHLLQVILINDNLLSSLPNLFRFVPLTHL DLRGNRIKLLPYVGLQHMDKVVELQLEEN PWNCSCELISLKDWLDSISYSALVGDVVCFT FRLHGRDLDEVKQELCPRRILSDYEMRPQTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVRPTSRQPSKDLGYSNYQPSIA YQTKSPVPLECPTACSCNLQISDLGLNVNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHLGNNIKSIMQDRAFGDLTN LRRLYLNGRRIELSFELFYGLQSLQYLFLQY NLIRRISGSTEDPVPNLQLLHLANNILLQAMPS GVFSG.TLTLRLNLRSNHFTSLPVSGVLDQLKS LIQDILHDNWDDCTCDIVVGMLWPEQLAVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTSSIQVPARTSAVTPAVRLINSTGA PASLAGGGGASVPLSVLLISLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSEN MQYSVYGGGGTGGHPHAHRGPALPK VKTPAGHVYEYPIPPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPP QQPQQPPPQLQLPPGLQCPCRESSHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLEFYFRFCSPAAYTFMYLGRPNLFW TPAGNSLEFYFRFCSPAAYTFMYLGRPN QYLHPGAGDSRLREPVLYSPPSAVFVEPNING TPAGNSLEFYFRFCSPAAYTFMYLGRPN TPAGNSLEFYFRFCSPAAYTFFNYLGRPN QYLHPGAGDSRLREPVLYSPPSAVFVEPNING YLLEKAKLNVEPDYLE LEKQTTFSQF SPSAAGGLAWVSLALGSGSRGNHGSGVGT AMAGALVRKAADYVRKDPRTYLLFACHATNEVA QULQGGRLIKHEMTKTASA QULQGGRLIKHEMTKTASA PTTTKLHHSSPLWNFFAQLL CYSLTFMRFAVKVQPRNWLLFACHATNEVA QULQGGRLIKHEMTKTASA PTTTKLHLSPSPLWNFFAQLL CYSLTFMRFAVKVQPRNWLLFACHATNEVA QULQGGRLIKHEMTKTASA PTTTKLLHSPSPLWNFFAQLL CYSLTFMRFAYKVQPRNWLLFACHATNEVA QULQGGRLIKHEMTKTASA PTTTKLLHSPSPLWNFFAQLL PSTATESTUCKRGNNAK PTTTKLLHSPSPLWNFFAQLL PSTATESTUCKRGNNAK PTTTKLLHSPSPLWNFFAQLL PSTATESTUCKRGNNAK PTTTKLLHSPSPLWNFFAQLL PSTATESTUCKRGNNAK PTTTKLLHSPSPLWNFFAQLL PSTATESTUCKRGNNAK PTTTKLLHSPSPLWNFFAQLL PSTATESTUCKRGNNAK PTTTKLLHSPSPLWNFFAQLL PSTATESTUCKRGNNAK PTTTKLLHSPSPLWNFFAQLL PSTATESTUCKRGNNAK PTTTKLLHSPSPLWNFFAQLL PSTATESTUCKRGNNAK PTTTKLLHSPS	1		ĺ				1
RTYYSSNYBACYEDMGNNTANWRMLIRIL POSSGETIPPLIMIFCYGETILTEKAHMGQK HRAMRVIFAVVLIFLLCWLPYNLVLLADTIM RTQVIQETCERRNHDRALDATELIGHISCLON PLIYAFIGQKFRIGLIKILAHGLISKDSLPKDS RPSFVGSSGHTSTTIL 1045 2395 A 8724 254 3184 FRANLATIYANRRGAQGCKMHTCCPPVTLEQ DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY YGEICDNACPCEEKDGILTYSCENRGIISLSEIS PPRFPIYHLLISGNILLNRI, YPNEFVNYTGASL HLGSNVIQDIETGAFHGLRGLRRI.HINNINL ELLRDDTFIGLENLEYLQWYNJISVIEPNAF GKLHLQVILINDNILSSLPNNLFRFVPLTHL DLRGNRIKLLPYVGLLQHMDKVVCETP FRI.HGRDLDEVSKQELCPRRIJSDYEMRPOTP LSTTGYLHTTPASVNSVASKYVEPPLKPP KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGINNOCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRIVINGNRIERISPELFYGLQSLQVIFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVPSGCTLLRINLRSNHFTSLPVSGVDQLKSV GVLVDEVICKAPKKFAETDMRSIKSEVDQLKSV ULVDEVICKAPKKFAETDMRSIKSELCPDYS DVVVSTPTPSQVPAATSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGIFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAVHHRGPALPK VKTPAGHVYEYIPHELGMCKNPIYRSREGN SVEDYKDLHELKYTYSSHLPDALPK VKTPAGHVYEYIPHELGMCKNPIYRSREGN SVEDYKDLHELKYTYSNHLLQQQQPPPP QQPQQPPPOLQLPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYFKFCSPAAYTTSRNYDLRRH QYLHPGAGDSRLREPVLYSPPSAVFENPNDLERRH QYLHPGAGDSRLREPVLYSPPSAVFENPNDLERRH QYLHPGAGDSRLREPVLYSPPSAVFENPNDR YLELKAKLNPEDYJELVEKQTTFSQP GPVANWGLPAAHDMKKSPEISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLKHEMTKTAS QLQGGRLKKHEMTKTAS PAKFFTSLQSLNWSSHLPPSFATESVGKRGNAK PPTTKLLINSSLWNFSHQVL							1
PQSSGIVPLLIMLFCYGFTLRTLFKAHMGQK HRAMRVIFAVVLIFLLCWLPYNLVLLADTLM RTQVIGETCERRNHIDRALDATELIGILHSCLN PLIYAFIGQKFRHGLLKILAHIGLISKDSLFKDS RPSFVGSSSGHTSTTL 1045 2395 A 8724 254 3184 FRANLAITVANRRGAQGGKMHTCCPPVTLEQ DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY YGEICDNACPCEEKDGILTYSCENRGIISLSEIS PREPIYHLLLSGNILLNRLYPNEFVNYTGASL HLGSNVQDIETGAFHGLRHLMNNKL ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF GKLHLQVLILNDNLLSSLENNLFREVPLTHL DLRGNRIKLKLPYVGLLQHMDKVVELQLEEN PWNCSCELISLKDWLDSISYSALVGDVCCTP FRHGRDLDEVSKQELCFRRLISDYEMRPOTP LSTIGYLHTIPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVRTISRQPSKDLGSVSNGDSIA YQTKSPPVJELCFTACSCNLQISDLGINVNCQE RKIESIABLQPKPYNRKMYLTENYIAVVRRT DLLEATGLDLHH,GNNRIGOPAFGDLTN LRRLYLNGRRIERLSPELFYGLQSLQYLFLQY NLIRBIQSGTTDFVPNLQLLFLINNLLQAMPS GVFSGITLIKLRISNHTSLPVSGVLDQLKS LQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGIFVLVMKRKKNQSHTSNSDVSSFN MQYSVYGGGGGTGGHPHAHVHRGPALPK VKTPAGHVYEYPPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQOPPPPP QQPQQOPPPQLQLQPGGEERRSHHLRSPAYS VSTEPPREDLISPVQDARFSQULFDCHHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRPH QYLHPGAGDSRLREPVLYSPPSAVFPPNRNE TPAGNSLPEYPKFPCSPAAYTFSPNYDLRPH QYLHPGAGDSRLREPVLYSPPSAVFYPPNRNE TPLELKAKLNVPDYLEUVLEKQTTFSQF VSELKAKLNVPDYLEUVLEKQTTFSQF PSPAANGLPLAANDMKKSPEISGRMTH-ALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTAS 1047 2397 A 8741 673 924 ALPOTPQTVTINTDGKVKSFTSFHSNPNLPP AKFFTSLQSLWWSHLPPSTATESVGKRGNAK PPTTKLLHSSPLWFFFATESVGKRGNAK PPTTTKLLHSSPLWFFFATESVGKRGNAK PPTTTKLHSSPLWFFFATESVGKRGNAK PPTTTKLHSSPLWFFFATESVGKRGNAK PPTTTKLHSSPLWFFFATESVGKRGNAK	1						
HRAMRVIFAVVLIFLICWLPYNLVLLADTI.M RTQVIQETCERRNHDRALDATEIL.GILHSCLN PLIYAFIGQKFRHGLLKILAHIGLISKDSLFKDS RPSFVGSSSGHTSTITL 1045 2395 A 8724 254 3184 FRANLAITVANRGAGGGKMHTCCPPVTLEQ DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY YGEICDNACPCEEKDGILTVSCENRGIISLSEIS PPRFPIYHLLLSGNLLNRLYPNEFVNYTGASIL HLGSNVIQDIETGAFHGLRGLRLHLNNNKL ELLRDDTFLGLENLEYLQVYNYISVIEPNAF GKLHLLQVLINDNLLSSLPNLFREFVLTHL DLRGNRLKLLPYVGLLQHMDKVVVEQLEEN PWNSCSCLISLKDWLDSISYSALVGDVVCETP FRLHGRDLDEVSKQELCPRRLISDYBMRPQTF LSTTGYLHTTPASVNSXATKYTEPVLTHL DLRGNRLKLLPYVGLLGLISNYRGPSKDLGYSNYGPSIA YQTKSPPVJECCPTACSCNLQISDLGJNNVQQE RKIESIAELQPKPYNPRKMTLTENYIAVVRRT DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLRSNHFTSLPVSGVLDQLKS LQDLHDNPWDCTCDIVGKKLWVEQUKVG VLVDEVICKAPKKFAETDMRSIKSSELCPDYS DVVVSTPTPSQVPAARTSAVTPAVRLNSTGA PASLGAGGGASSVLSVLILSLLLFMSVVA AGIFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGOGGGTGGPHAHVUHRRGPALPK VKTPAGHVYEVIPHPLGHMCKNPITYSRGEON SVEDYKDLHELKVTYSSNHLQQQQQPPPPP QQPQQQPPPQLQLQPGGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYKFFCSSPAAYTSSNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEULEKQTITSQF JPLEKAKLNVEPDYLEULEKQTITSQF GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFRAYKVQPRNWLLFACHATNEVA QLIQGGRLMVSSTLAGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFRAYKVQPRNWLLFACHATNEVA QLIQGGRLMKHEMTKTAA QLIQGGRLMKHEMTKTAA QLIQGGRLMKHEMTKTAA	1			[
RTQVIQETCERRNHIDRALDATELIGILHSCLN PLYAFIGQKFRHGLKILAHGLISKDSLPKDS RPSPVGSSSGHTSTIL 1045 2395 A 8724 254 3184 FRANLATIVANRGGAGGKMHTCCPPVTLEG DLHRKMHSWM_QTLAFAYTSLVLSCAETIDY YGEICDNACPCEEKDGILTVSCENRGIISLSEIS PPRFPIYHLLISGNLLNRLYPNEVNYTGASIL HLGSNVIQDIETGAFFGLGGLRRLHHNNNKL ELLRDDTFLGLENLEYLQVDYNYISVIEDNAF GKLHLLQVLILNDNLLSLPNNLFREPVLTHL DLRGNRLKLLPYVGLLQHMDKVVELQLEEN PWNCSCELISLKDWLDISISYALVGDVVCETF FRHGRODLDEVSKQELGPRALISDVERRPOTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGINNNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLHLGNNSMIGDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLQAMPS GVFSGLTILIRINLRSNHFTSLPVSGVLDQLKS LIQDLHDNPWDCTCDIVGMKLWYEQUKVG YLVDEVICKAPKKFAETMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLLILLLYFIMSVYVA AGGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHRGPALPK VKTPAGHVYEYIPPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQQPGEERRESHHLRSPAYS VSTIEPREDLLSPYQDAFYRGILEPBKHCST TPAGNSLEYFKFCSPAAYTESPNYDLRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDVLEVLEQTITSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVELVLEKQTITSQF ULELKAKLNVEPDVLEVLEKQTITSQF THEAGNSLEYFKFCSPAAYTESOYLMSTHFW GPVANWGLPIAANIDMKKSPELISGRMTFALC CYSLTFMRFAYKVQPRNVLLPACHATNEVA QLIQGGGLIKHEMTKTASA 1047 2397 A 8741 673 924 ALFGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFTSLQSLINKESHLPPSATESVGKRGNAK PPTTKLLHSSPLNMFSAQQL			ļ				
PLIVAFIGQKFRHGLIKILAHGLISKDSLPKDS RPSPVGSSSGHTSTTL 1045 2395 A 8724 254 3184 FRANLAITVANRRGAQGGKMHTCCPPVTLEQ DLHRKMHSWMLQTLAFAYTSLVLSCAETIDY YGEICDNACPCEERDGILTYSCENRGIISLSEIS PRFPIYHLLLSGNLINRLYPNEFVNYTGASIL HLGSNVIQDIETGAFHGLIGLIKHLNNNKL ELKRDDTFLGLENLEYLQVDYNYISVEDNAF GKLHLLQVLILNDNLLSSLPNALFREVPLTHL DLRGNRLKLLPYVGLLQHIMDKVVELQLEEN PWNCSCELISLKDWLDSISYSALVGDVVCETP FRLHGRDLDEVSKQELCPRILISDYEMRPQTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVPFTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDIGLNVNCQE RKIESIAELQPKPYPNFKMYLTENYIAVVRRT DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLGSLQYLFLQY NLIREIQSGTFDPYPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKPAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPSAVRLNSTGA PASLGAGGASSVPLSVLILSLLLVFIMSVVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHRGPALPK VKTPAGHVYEVIPPLSVSLLLSLLLVFINSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHRGPALPK VKTPAGHVYEVIPPLAGOSNLERPH QQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEFREDLLSPVQDADRFYRGILEFDKHCST TPAGNSLPEYPKFPCSPAAYTFSRYDLRRH QYLLHFAGOSNLERPYLSPSAVTSSPNYDLRRH QYLHPGAGOSNLERPYLSPSAVTSPNYDLRRPH QYLHPGAGOSNLERPYLSPSPAVTSPPRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGROBISGSGVGT AMAGALVRKAADYVSLPPLACHATNEVA QLIQGGGLIKHEMTKASA CLQGGGRIKHEMTKASA ALPGTPQQTVTLINTDGKVKSFTSPHSNPNLPP AKFTSLQSLINKENSHLPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL			1				l =
RPSPVGSSSGHTSTIL 1045 2395 A 8724 254 3184 FRANLATIVANRRGAQGKMHTCCPPVTLEQ DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY YOEICDNACPCEERDGILTYSCENRGIISLSES PREPTIFIHLLSGNILDRILYPNEFVNYTGASIL HLGSNVIQDIETGAFHGLRGLRRLHLNNNKL ELLRDDTFLGLENLEYLQVDYNTSVIEPNAF GKLHLLQVLILNDNILLSLPNNLFREPVDLTHL DLRGNRLKLLPPVGLLQHMDKVVELQLEEN PWNCSCELISLKDWLDISSYSALVGDVVCETP FRLHGRDLDEVSKQELCPRRLISDVEMRPQTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNRPVRPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGLNNNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHGNNRISMIGDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLREIGSGTFDPYPNL DLFLENNILLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKRAENMSISSALCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHRGPALPK VKTPAGHVYEYPPHLGMCKNPTYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQPGEERRESHHLRSPAYS VSTIEPREDLLSPYQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTESPNYDLRRH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDVLEVLEKQTITSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGDHSGSOVGT AMAGALVKLAADVYSKDPRDYLMSTHFW GPVANWGLPIAANDMKKSPEUSGRMTFALC CYSLTFMRFAYKVQPRNWLFACHATNEVA QLQGGGLIKHEMTKTANLSA LQQGGRIKHEMTKTANLSA LQQGGRIKHEMTKTANLSA LQQGGRIKHEMTKTANLSA			1	(l		_ ^
1045 2395 A 8724 254 3184 FRANLATTVANRRGAQGKMHTCCPPVTLEO DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY YGEICDNACPCEEKDGILTVSCENRGIISLSEIS PPRFPIYHLLISGNLLNRLYPNEFVNYTGASIL HLGSNVIQDIETGAFHGLIRGLRRLHLNNNKL ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF GKLHLQVLLINDNLLSSLPNNLFRFVPLTHL DLRGNRIKLLPYVLQLHDMB, VVELQLEEN PWNCSCELISLKDWLDSISYSALVGDVVCETP FRLHGRDLDEVSKQELCPRRLISDYEMRRQTIP LSTTGYJHTTPASVXATSSSAVYKPPLKPP KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGLNNNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLIHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPEJFYGLQSLQYLFLQY NLIREIQSGTFDYPNLQLLFINNNLQAMPS GVFSGLTLRLNLRSNHFTSLPVSGVLDQLKS LIQUDLHDNPWDCTOLYGMKLWVEQULKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVSLYSLISLLLLYEMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYJPHPLGHMCKNPTYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPQOLQOQPPPPP QQPQQPQPPQLOPQGPPPOLQPPPPP QQPQPPQDADRYPRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREVPLYSPPSAVYEPPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSGRGDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLJNSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRAYKVQPRNWLFACHATNEVA QLIQGGRIKHEMTKTASA QLIQGGRIKHEMTKTASA QLIQGGRIKHEMTKTASA PTTKLLCACYSLTFTSPNYDLRPH AKFFTSLQSLNWSSHLIPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL							
DLHRKMHSWMLQTLAFAVTSLVLSCAETIDY YGEICDNACPCEEKDGILTYSCENRGISLSEIS PREFFIYHLLLSGNLNRLYPNEFVNTGASIL HLGSNVIQDIETGAFHGLRGLRRLHLNNNKL ELLRDDTFLGLENLEYLQVUDYNTISVEPPNAF GKLHLLQVLLINDNLLSSLPNNLFRFVPLTHL DLRGNRLKLLPYVGLLQHMDKVVELQLEEN PWNCSCELISLKDWLDSISYSALVGDVVCETP FRLHGRDLDEVSKQELCPRRLISDYEMRPQTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSLA YQTKSSVPLECPTACSCNLQISDLGLNVNCQE RKIESIAELQPRPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHLGINNISMIGDRAFGDLTN LRRLYLNGRNERLSPELFYGLQSLQYLFLQY NLIREIQSGTTDPVPNLQLLFLNNNLQAMPS GVFSGLTLRLNLRSHTFSLPVSGVLDQLKS GVFSGLTLRLNLRSHTFSLPVSGVLDQLKS LIQDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVYSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRKKNQSDHTSNNSDVSSFN MQYSVYGGGGTGGHPHAHVHIRGPALPK VXTPAGHVYSYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPYPKFPCSPAAYTFSPNYDLRPH QVLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSOF TPAGNSLPLYPKFPCSPAAYTFSPNTDLRPH QVLHPGAGDSRLREPVLYSPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSOF TPAGNSLPLYPKFPCSPAAYTFSPNTDLRPH QVLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSOF TPAGNSLPLYFKFFCSPAAYTFSPNTDLRPH QVLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSOF TPAGNSLPLYFKFFCSPAAYTFSPNTDLRPH QVLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKGTTFSOF TPAGNSLPYKFFCSPAAYTFSNFNLTPH QPVANWGLPIAANDMKKSPEISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTYTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLIPSPAATESVQKRGNAK PPTTKLLHSSPLWNFFAQQL	1045	2305	_	8724	254	2194	
YGEICDNACPCEEKDGILTVSCENRGIISLSEIS PPRFPIYHLLLSGNLLNRLYPNEFVNYTGASIL HLGSNVIQDIETGAFHGLRGLRRLHINNIKL ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF GKLHLLQVLLNDNLLSSLPNNLFRFVPLTHL DLRGNRLKLPYVGLQHMDKVVELQLEEN PWNCSCELISLKDWLDSISYSALVGDVVCETP FRLHGRDLDEVSKQELCPRRLISDYBMRPQTP LSTTGYLHITPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA YQTKSSPVPLECPTACSCPSKDLGYSNYGPSIA YQTKSSPVPLECPTACSCPSKDLGYSNYGPSIA YQTKSSPVPLECPTACSCPSKDLGYSNYGPSIA YQTKSSPVPLECPTACSCNLQISDLGINNICQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLHLGNNRISMIQDRAFGDLTN LRRLYLGNGRIERLSPLFYGLQSLQYLFLQY NLREIQSGTTDPVPNLQLLFLNNNLLQAMPS GVFSGLTLRLNLRSNHFTSLPVSGVLDQLKS LIQDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLLISLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGGTGPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDVKDLHELKVTYSSNHHLQQQQQPPPPP QQPQPPPQLQLQPGEEERRSHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTTSPNYDLRRPH QVLHPGAGDSRLREPVLYSPPSAUFVEPNNBL YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSGRDHSGSGVGT AMAGALVRKAADVYRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIJSGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPSATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	1043	2373	^	0/24	234	3104	l = .
PPRFIYILLISGNILNRI,YPNEFVNYTGASIL HLGSNVIQDIETGAFHGLRGLRILHINNNKL ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF GKLHILQVILINDNILLSSLPNT,FRFYPLTHIL DLRGNRLKKLPYVGLLQHMDKVVELQLEEN PWNCSCELISLKOWLDSISYSALVGDVVCETP FRIHGRDLDEVSKQELCPRIJSDYEMRPQTP LSTTGYLHTIPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVYRYTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGLNVNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLIHGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY MLIREIQSGTFDPYPNLQLIFINNNLLQAMPS GVPSGLTLIKINLRSNHFTSLPVSGVLDQLKS LIQDIHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFINSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDVKDLHELKVTYSSNHILQQQQQPPPPP QQPQQPPPQLQLQPGEEERESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMFFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPSPATESVGKRGNAK PPTTKLLHSSEL WNFFAQQL 1046 1047 1047 1047 1047 1047 1047 1047 1047							
HLGSNYQDIETGAFHGLRGLRRLHLNNNKL ELLRDDTFLGLENLEYLQVDYNYISVIEPNAF GKLHLLQVILINDNILLSSLPNLFRFYPLTHL DLRGNRLKLLPYVGLQHMDKVVELQLEEN PWNCSCELISLKDWLDSISYSALVGDVVCETP FRLHGROLDEVSKQELCPRRLISDYEMRPQTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGLNVNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLMGNRLESPELFYGLOSLQVLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLQAMPS GVPSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDVS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHIRGPALPK VKTPAGHVYEYPIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLVSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLJAANDMKKSPELISGRMTFALC CYSLTMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQIVTLNTDGKVKSFTSPHSNFNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLPVSPATESVGKRGNAK PPTTKLLHSSPLPVSPATESVGKRGNAK	l i						
ELLRDITEJENLEYJQVDYNYISVIEPNAF GKLLQVLILNDNLLSSLPNNLFFFVPLTHL DLRGNRIKLLPYYGLLQHMDKVVELQLEEN PWNCSCELISLKDWLDSISYSALVGDVVCETP FRLHGRDLDEVSKQELCPRRLISDYEMRPQTP FRLHGRDLDEVSKQELCPRRLISDYEMRPQTP LSTTGYLHITTPASVNSVATSSAYYKPPLKPP KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGINVNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLGSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDVS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLLISLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGPPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHEKVTYSSNHHLQQQQPPPP QQPQQPPPQLQLQPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFFCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTITFSQF SPSAAGGLAWVSLALGSGSGRGDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAXVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPITKLLHSSPLWNFFAQQL							
GKLHLLQVLILNDNLLSSLPNLT,FRFVPLTHIL DLRGRIKLLPYVGLLQHMDKVVELQLEEN PWNCSCELISK, LOWLDSISYSAL VGDVVCETP FRI.HGRDLDEVSKQELCPRRLISDYEMRPQTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGLNVNCQE RKIESIAELQPKPYNFKKMYLTENYIAVVRRT DLLEATGLDLHLGNNRISMIODRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLTREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNESNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDVS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLLISLLLVFIMSVFVA AGLFVLVMKRRKKNOSDHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHHRGPALPK VKTPAGHVVEYIPHPLGHMCKNPTYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQQPPPPP QQPQQQPPPQLQLQPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPLAAINDMKKSPEIJSGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL							
DLRGNRLKLLPYVGLLQHMDKVVELQLEEN PWNCSCELISLKDWLDSISYSALVGDVVCETP FREHGRDLDEVSKQELCPRRLISDYEMRPQTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVRPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGLNVNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGRNIERLSPELFYGLGSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQOPPPPP QQPQQQPPPQLQLQPGEERRESHHLRSPAYS VSTIEPREDLLSVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGDHSGSGVGT AMGALVRKAADYVSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	i i		'				
PWNCSCELISLKDWLDSISYSALVGDVVCETP FRLHGRDLDEVSKQELCPRRLISDYEMRPQTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGLNVNCQE RKIESIAELQPKPYNPKKMYLTENYLAVVRRT DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVNSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGIMCKNPTYRSREGN SVEDVKDLHELKVTYSSNHHLQQQQPPPPPP QQPQQPPPQLQLQPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVEKQTTFSQF SPSAAGGLAWVSLALGSGSRGDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRRAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVILNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL							DI DENDI KI I DVVCI I OUMDKAVELOI EEN
FRLHGRDLDEVSKQELCPRRLISDYEMRPQTP LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRYPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGLNVNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHIGNNRISMIQDRAFGDLTN LRRLYLNGRRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRILRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLLISLLLVFIMSVFVA AGLFYLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGPHAHVHHRGPALPK VKTPAGHVYEYPPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQQPPPQLQPGEERRSSHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLEEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLTNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL							
LSTTGYLHTTPASVNSVATSSSAVYKPPLKPP KGTRQPNKPRVRRTSRQPSKDLGYSNYGPSIA YQTKSPVPLECPTACSCNLQISDLGLNVNCQE RKIESIAELQPKPYNPKKMYLIENYIAVVRRT DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLR.NLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTINTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	1 1			1			
KGTRQPNKPRVRTPTSRQPSKDLGYSNYGPSIA YQTKSPVPLECTTACSCNLQISDLGLNVNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGGSSVPLSVLLSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPLYSPSAVFVEPNRNE QYLHPGAGDSRLREPVLYSPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVKAADYVRSKDFRDYLMSTHHW GPVANWGLPIAANDMKKSPEIJSGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVILNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	1 1						
YQTKSPVPLECPTACSCNLQISDLGLNVNCQE RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQQPPPPP QQPQQPPPQLQLQPGEEERRSHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTTSPNYDLRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSFEIJSGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL		· .					
RKIESIAELQPKPYNPKKMYLTENYIAVVRRT DLLEATGLDLLHLGNNRISMIODRAFGDLTN LRRLYLNGNRIERLSPELFYGLGSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPIPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQLQPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRIKHEMTKASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLOSLNWSSHLPPSFATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL							
DLLEATGLDLLHLGNNRISMIQDRAFGDLTN LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNIHLQQQQPPPPP QQPQQQPPPQLQLQPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTITFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLOSLNWSSHLPPSFATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL							
LRRLYLNGNRIERLSPELFYGLQSLQYLFLQY NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRLNRSNHFTSPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	1	' I					· _ · _ ·
NLIREIQSGTFDPVPNLQLLFLNNNLLQAMPS GVFSGLTLLRINLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHILQQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL							
GVFSGLTLLRLNLRSNHFTSLPVSGVLDQLKS LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQQPPPPP QQPQQPPPQLQLPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSFATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL							
LIQIDLHDNPWDCTCDIVGMKLWVEQLKVG VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQLQPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	i i			1			
VLVDEVICKAPKKFAETDMRSIKSELLCPDYS DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNIHLLQQQQPPPPP QOPQQOPPPQLQLOPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL					ļ	İ	
DVVVSTPTPSSIQVPARTSAVTPAVRLNSTGA PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNIHLQQQQPPPPP QQPQQQPPPQLQLQPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL]						
PASLGAGGGASSVPLSVLILSLLLVFIMSVFVA AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQQPPPQLQLQPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	1 1			ļ			
AGLFVLVMKRRKKNQSDHTSTNNSDVSSFN MQYSVYGGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNHHLQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL							
MQYSVYGGGGTGGHPHAHVHHRGPALPK VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNIHLQQQQPPPPP QQPQQPPPQLQLQPGEEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	[- 1		1	1	i	
VKTPAGHVYEYIPHPLGHMCKNPIYRSREGN SVEDYKDLHELKVTYSSNIHLQQQQPPPPP QQPQQPPPQLQLQPGEERRESHILRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	1 .		1				
SVEDYKDLHELKVTYSSNIHLQQQQPPPPP QQPQQPPPQLQLQPGEERRESHHLRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	-						
QQPQQPPPQLQLQPGEERRESHHILRSPAYS VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	[İ	1	ŀ	
VSTIEPREDLLSPVQDADRFYRGILEPDKHCST TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL		İ					
TPAGNSLPEYPKFPCSPAAYTFSPNYDLRRPH QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAANDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL		į		. 1			
QYLHPGAGDSRLREPVLYSPPSAVFVEPNRNE YLELKAKLNVEPDYLEVLEKQTTFSQF 1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL		1		ł	l	l	TDAGNGI PEVPKEDCEDA AVTECDATVIT DEDIS
1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	1 1			٠		ļ	OVI HEGA GUSEI BEDVI VEDDE V VEVEDNOVIE
1046 2396 A 8736 28 452 SPSAAGGLAWVSLALGSGSRGRDHSGSGVGT AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL			1		Ì		
AMAGALVRKAADYVRSKDFRDYLMSTHFW GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	1046	2396	$\overline{\mathbf{A}}$	8736	28	452	
GPVANWGLPIAAINDMKKSPEIISGRMTFALC CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	ا ```` ا		^	3730	20	734	
CYSLTFMRFAYKVQPRNWLLFACHATNEVA QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	[]		- 1	l			
QLIQGGRLIKHEMTKTASA 1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	}			!			
1047 2397 A 8741 673 924 ALPGTPQQTVTLNTDGKVKSFTSPHSNPNLPP AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL		ļ		Ì			
AKFFTSLQSLNWSSHLPPSPATESVGKRGNAK PPTTKLLHSSPLWNFFAQQL	1017	2207		0741	673	<u> </u>	
PPTTKLLHSSPLWNFFAQQL	104/	7371	^	0/41	0/3	724	
]		ļ			-
1046 2398 A 8/4/ 3 3054 PEVTKPSLSQPTAASPIGSSPSPPVNGGNNAKR	1040	2200	<u>,</u>	9747	<u> </u>	5054	
	1048	2398	A	5/4/	3	5054	PEVIKPSLSQPTAASPIGSSPSPPVNGGNNAKR

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon.
				peptide sequence		/=possible nucleotide deletion, \-possible nucleotide insertion
				•		VAVPNGQPPSAARYMPREVPPRFRCQQDHK VLLKRGQPPPPSCMLLGGGAGPPPCTAPGAN
						PNNAQVTGALLQSESGTAPDSTLGGAAASNY ANSTWGSGASSNNGTSPNPIHIWDKVIVDGS
						DMEEWPCIASKDTESSSENTTDNNSASNPGSE KSTLPGSTTSNKGKGSQCQSASSGNECNLGV
					'	WKSDPKAKSVQSSNSTTENNNGLGNWRNVS
						GQDRIGPGSGFSNPNPNSNPSAWPALVQEGTS RKGALETDNSNSSAQVSTVGQTSREQQSKME
						NAGVNFVVSGREQAQIHNTDGPKNGNTNSL NLSSPNPMENKGMPFGMGLGNTSRSTDAPSQ
						STGDRKTGSVGSWGAARGPSGTDTVSGQSNS GNNGNNGKEREDSWKGASVQKSTGSKNDS
						WDNNNRSTGGSWNFGPQDSNDNKWGEGNK
						MTSGVSQGEWKQPTGSDELKIGEWSGPNQPN SSTGAWDNQKGHPLLENQGNAQAPCWGRSS
						SSTGSEVEGQSTGSNHKAGSSDSHNSGRRSY RPTHPDCQAVLQTLLSRTDLDPRVLSNTGWG
						QTQIKQDTVWDIEEVPRPEGKSDKGTEGWES AATQTKNSGGWGDAPSQSNQMKSGWGELS
						ASTEWKDPKNTGGWNDYKNNNSSNWGGGR
						PDEKTPSSWNENPSKDQGWGGGRQPNQGWS SGKNGWGEEVDQTKNSNWESSASKPVSGWG
						EGGQNEIGTWGNGGNASLASKGGWEDCKRS PAWNETGRQPNSWNKQHQQQPPQQPPPPQ
						PEASGSWGGPPPPPPGNVRPSNSSWSSGPQPA
				:		TPKDEEPSGWEEPSPQSISRKMDIDDGTSAWG DPNSYNYKNVNLWDKNSQGGPAPREPNLPTP
				•		MTSKSASDSKSMQDGWGESDGPVTGARHPS WEEEEDGGVWNTTGSQGSASSHNSASWGQG
					,	GKKQMKCSLKGGNNDSWMNPLAKQFSNMG LLSQTEDNPSSKMDLSVGSLSDKKFDVDKRA
			ľ	İ		MNLGDFNDIMRKDRSGFRPPNSKDMGTTDS
						GPYFEKGGSHGLFGNSTAQSRGLHTPVQPLN SSPSLRAQVPPQFISPQVSASMLKQFPNSGLSP
						GLFNVGPQLSPQQIAMLSQLPQIPQFQLACQL LLQQQQQQLLQNQRKISQAVRQQQEQQLA
					ĺ	RMVSALQQQQQQQQQRQPGMKHSPSHPVGPK
						PHLDNMVPNALNVGLPDLQTKGPIPGYGSGF SSGGMDYGMVGGKEAGTESRFKQWTSMME
			1			GLPSVATQEANMHKNGAIVAPGKTRGGSPY NQFDIIPGDTLGGHTGPAGDSWLPAKSPPTNK
			Ī			IGSKSSNASWPPEFQPGVPWKGIQNIDPESDP
				1		YVTPGSVLGGTATSPIVDTDHQLLRDNTTGS NSSLNTSLPSPGAWPYSASDNSFTNVHSTSAK
					İ	FPDYKSTWSPDPIGHNPTHLSNKMWKNHISS RNTTPLPRPPPGLTNPKPSSPWSSTAPRSVRG
		İ		1	1	WGTQDSRLASASTWSDGGSVRPSYWLVLHN LTPQIDGSTLRTICMOHGPLLTFHLNI,TOGTA
Ì			į			LIRYSTKQEAAKAQTALHMCVLGNTTILAEF
					i e	ATDDEVSRFLAQAQPPTPAATPSAPAAGWQS LETGQNQSDPVGPALNLFGGSTGLGQWSSSA
						GGSSGADLAGASLWGPPNYSSSLWGVPTVED PHRMGSPAPLLPGDLLGGGSDSI
1049	2399	A	8748	200	1387	VPWKRQDEQLSLQVETLYLDSPAVIHLLSPTF
						LPPSSLPPFLQIVDSSSSACTLDSFFPFLAPWDS PQDCGFKDHQPLTLQALTVELARWTLMLLLS
		l			ļ	TAMYGAHAPLLALCHVDGRVPFRPSSAVLLT ELTKLLLCAFSLLVGWQAWPQGPPPWRQAA
			1			PFALSALLYGANNNLVTYLQRYMDPSTYQVL

<u> </u>	L OFO ID	1 3 (04	000	Deading 1	15	
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1	in USSN	nucleotide location	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq- uence		09/496	correspondi	corresponding to last amino	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
seq- uence	ucnce	}	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
uciice	}	J	314	amino acid	of peptide	
i				residue of	sequence	T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon,
		1		peptide	sequence	/=possible nucleotide deletion, \=possible
1	1	l		sequence		nucleotide insertion
 	ļ <u></u> .	 	 	Sequence	ļ	
· .				•		SNLKIGSTAVLYCLCLRHRLSVRQGLALLLL
i		1	1			MAAGACYAAGGLQVPGNTLPSPPPAAAASP
					Ì	MPLHITPLGLLLLILYCLISGLSSVYTELLMKR
		l				QRLPLALQNLFLYTFGVLLNLGLHAGGGSGP
					1	GLLEGFSGWAALVVLSQALNGLLMSAVMKH GSSITRLFVVSCSLVVNAVLSAVLLRLQLTAA
[ĺ			ĺ	
1050	2400	A	8758	3	1660	FFLATLLIGLAMRLYYGSR
1030	2400	^	0/30	3	1000	WVSSMGFEELLEQVGGFGPFQLRNVALLALP
						RVLLPLHFLLPIFLAAVPAHRCALPGAPANFS
					ļ	HQDVWLEAHLPREPDGTLSSCLRFAYPQALP
						NTTLGEERQSRGELEDEPATVPCSQGWEYDH
						SEFSSTIATESQWDLVCEQKGLNRAASTFFFA
						GVLVGAVAFGYLSDRFGRRRLLLVAYVSTLV
			1.	•		LGLASAASVSYVMFAITRTLTGSALAGFTIIV
						MPLELEWLDVEHRTVAGVLSSTFWTGGVML
		l	1			LALVGYLIRDWRWLLLAVTLPCAPGILSLWW
		Ì				VPESARWLLTQGHVKEAHRYLLHCARLNGR
		l				PVCEDSFSQEAVSKVAAGERVVRRPSYLDLF
		1		1		RTPRLRHISLCCVVVWFGVNFSYYGLSLDVS
						GLGLNVYQTQLLFGAVELPSKLLVYLSVRYA
		-				GRRLTQAGTLLGTALAFGTRLLVSSDMKSWS
				į į		TVLAVMGKAFSEAAFTTAYLFTSELYPTVLR
		}				QTGMGLTALVGRLGGSLAPLAALLDGVWLS
		i				LPKLTYGGIALLAAGTALLLPETRQAQLPETI
1051	2401	A	8759	515	1625	QDVERKSAPTSLQEEEMPMKQVQN
1031	2401	Α .	0/39	313 ,	1023	EIRTPVAVSSAPSGDSEGDEETTQDEVSSHTS
						EEDGGVVKVEKELENTEQPVGGNEVVEHEV
				. *		TGNLNSDPLLELCQCPLCQLDCGSREQLIAHV
						YQHTAAVVSAKSYMCPVCGRALSSPGSLGR
						HLLIHSEDQRSNCAVCGARFTSHATFNSEKLP EVLNMESLPTVHNEGPSSAEGKDIAFSPPVYP
						AGILLVCNNCAAYRKLLEAQTPSVRKWALRR
			Ì			QNEPLEVRLQRLERERTAKKSRRDNETPEERE
					•	VRRMRDREAKRLORMOETDEORARRLORDR
						EAMRLKRANETPEKROARLIREREAKRLKRR
						LEKMDMMLRAOFGODPSAMAALAAEMNFF
						OLPVSGVELDSOLLGKMAFEEONSSSLH
1052	2402	A	8763	1106	70	RHGHGGRDRRGGGRVARPGGLGRYPGRGAA
		**	3,03	-100	, •	ASLVFVPTRRRSGPSGTASVAAMAYHSGYGA
						HGSKHRARAAPDPPPLFDDTSGGYSSOPGGY
						PATGADVAFSVNHLLGDPMANVAMAYGSSI
						ASHGKDMVHKELHRFVSVSKLKYFFAVDTA
						YVAKKLGLLVFPYTHONWEVOYSRDAPLPP
		.	<u>,</u>		_	RODLNAPDLYIPTMAFITYVLLAGMALGIQK
1				' I	1	RFSPEVLGLCASTALVWVVMEVLALLLGLYL
				l		ATVRSDLSTFHLLAYSGYKYVGMILSVLTGL
1				i		LFGSDGYYVALAWTSSALMYFIVRSLRTAAL GPDSMGGPVPRORLOLYLTLGAAAFOPLIIY
1				1	l	
1053	2403	A	8768	2	712	PRED VILVER DEL CA A APRIL CHETA DOLLAR
,000	2403	^	0/00	-	112	RPPRVWYPELRELSAAAPRWSHRTAPGIMVF
1					l	YFTSSSVNSSAYTIYMGKDKYENEDLIKHGW
					İ	PEDIWFHVDKLSSAHVYLRLHKGENIEDIPKE
				1		VLMDCAHLVKANSIQGCKMNNVNVVYTPW
ſ				ſ	ļ	SNLKKTADMDVGQIGFHRQKDVKIVTVEKK VNEILNRLEKTKVERFPDLAAEKECRDREER
l				ļ		
j				ļ	ļ	NEKKAQIQEMKKREKEEMKKKREMDELRSY
1054	2404	A	8769	344	527	SSLMKVENMSSNQDGNDSDEFM REATTLACRNSCWVFSRCSLGACKPTVCSMP
.054	2701	^	0/07	244	341	SLSRQGSQTLCLRLAEYCMESVDSQRLLLS
	1					3に3パグロ3グ1ドゲドば下せた I CWE3 ∧ D3グ以下下7

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
						QQESPAAGAARMNCKEGTDSSCGCRGNDEK KMLKCVVVGDGAVGKTCLLMSYANDAFPEE YVPTVFDHYAVTVTVGGKQHLLGLYDTAGQ EDYNQLRPLSYPNTDVFLICFSVVNPASYHNV QEEWVPELKDCMPHVPYVLIGTQIDLRDDPK TLARLLYMKEKPLTYEHGVKLAKAIGAQCYL ECSALTQKGLKAVFDEAILTIFHPKKKKKRCS EGHSCCSII
1056	2406	Α	8773	261	332	NPRIQLSGNSCCAGSCRVWLSEQ
1057	2407	A	8778	3	477	PAGIRHEQARGADRMGKCRGI.RTARKLRSH RRDQKWHDKQYKKAHLGTALKANPFGGAS HAKGIVLEKVGVEAKQPNSAIRKCVRVQLIK NGKKITAFVPNDGCLNFIEENDEVLVAGFGR KGHAVGDIPGVRFKVVKVANVSLLALYKGK KERPRS
1058	2408	A	8808	171	881	PGLSQEPSGSMETVVIVAIGVLATIFLASFAAL VLVCRQRYCRPRDLLQRYDSKPIVDLIGAME TQSEPSELELDDVVITNPHIEAILENEDWIEDA SGLMSHCIAILKICHTLTEKLVAMTMGSGAK MKTSASVSDIIVVAKRISPRVDDVVKSMYPPL DPKLLDARTTALLLSVSHLVLVTRNACHLTG GLDWIDQSLSAAEEHLEVLREAALASEPDKG LPGPEGFLQEQSAI
1059	2409	A	8809	246	757	MRLQGAIFVLLPHLGPILVWLFTRDHMSGWC EGPRMLSWCPFYKVLLLVQTAIYSVVGYASY LVWKDLGGGLGWPLALPLGLYAVQLTISWT VLVLFFTVHNPGLALLHLLLLYGLVVSTALI WHPINKLAALLLLPYLAWLTVTSALTYHLWR DSLCPVHQPQPTEKSD
1060	2410	A	8810	304	381	PKLSVYPLQSHHCLSEPFQSLVCCLA
1061	2411	A	8820	1673	848	SCKTENLLEMWWFQQGLSFLPSALVIWTSAA FIFSYITAVTLHHIDPALPYISDTGTVAPEKCLF GAMLNIAAVLCIATTYVRYKQVHALSPEENVI IKLNKAGLVLGILSCLGLSIVANFQKTTLFAA HVSGAVLTFGMGSLYMFVQTILSYQMQPKIH GKQVFWIRLLLVIWCGVSALSMLTCSSVLHS GNFGTDLEQKLHWNPEDKGYVLHMITTAAE WSMSFSFFGFFLTYIRDFQKISLRVEANLHGL TLYDTAPCPINNERTRLLSRDI
1062	2412	A	8824		763	GGAPPASVPARESPVSGAQGSSRTRGHKRAA GARAPQLCSSWQRRSAPAMSRGLQLLLLSCA YSLAPATPEVKVACSEDVDLPCTAPWDPQVP YTVSWVKLLEGGEERMETPQEDHLRGQHYH QKGQNGSFDAPNERPYSLKIRNTTSCNSGTYR CTLQDPDGQRNLSGKVILRVTGCPAQRKEET FKKYRAEIVLLALVIFYLTLIIFTCKFARLQSI FPDFSKAGMERAFLPVTSPNKHLGLVTPHKT ELV
1063	2413	A	8826	147	627	CETSTSSAGHAPCRHAAQGPPAEPTGLRLCSE HQRLHAWPPGPRRPSLWPPKNGKWHSGKRT AGGRPQRRPSRRQSQRPSAWSGSPRMHSPGQ KCSLMCPHRSQDSLSTAIFQRSPGANTGRALH CVLSKEMKSVQRSLGLSRIHLQSKRKIIHFVL TR
1064	2414	A	8835	2982	1869	LKDTLKSOMTOEASDEAEDMKEAMNRMIDE LNKQVSELSQLYKEAQAELEDYRKRKSLEDV TAEYIHKAEHEKLMQLTNVSRAKAEDALSE MKSQYSKVLNELTOLKQLVDAQKENSVSITE HLQVITTLRTAAKEMEEKISNLKEHLASKEVE

SEQ ID NO: of much colide sequence (A-Alainne C-Cysteine, peptide sequence (A-Alainne		T	1				
mucle cettide seg-	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
enice being seguence Seguence		1	noa	1			
Sequence	1		1				
1066	1		i				
amino acid residue of peptide residue of peptide residue of peptide sequence peptide sequence peptide sequence per sequenc		uence	1				
residue of peptide pep	uence		1	914			Q=Glutamine, R=Arginine, S=Serine,
Peptide		}	1	Ì	4		
		1	l			sequence	
VAKLEKQLEEKAAMTDAMYPRSSYEKIQS SLSESVALASKLKESVEKEKHYBEVVQIRS SLSESVALASKLKESVEKEKHYBEVVQIRS SLSESVALASKLKESVEKEKHYBEVVQIRS EVSQVKREKENIQTILLKSKEQEVQELLQKFQ QAQEELAEMKRYSESSKLEEDKAKKHEMS KEVTKLKEALNSLQLSYSTSSKRQQQLEA LQQVKQLQQMAEDVQKVQLQQHCEVISVYRMHL LYAVQQQMDEDVQKVLQQITMCKNQQK K		ļ	ł			ĺ	
SLESEVSVLASKILKESVEKEKEVTSELVQIPK EVSQVKREKENIQITLKSKEGOVNELLQKFQ QAQEELAEMKRYSESSSKLEEDKDKKINEMS KEVTKLKEALNISLGSLSYSTSSKRSQQLEA LQQVKQLQNQLAECKKQHGEVISVYRMHL LYAVQGQMDEDVQKVLKQILTMCKNQSQK K			<u> </u>		sequence		
BEVSQVKREKENIOTILLKSKEGEVNELLQKPG			}			1	
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,			1		l		
		1	ļ	ļ			EVSQVKREKENIQTLLKSKEQEVNELLQKFQ
LQQQVKQLQNGLAECKKQHQEVISVYRMHL LYAVQGQMDEDVQKVLKQILTMCKNQSQK K			i	İ			QAQEELAEMKRYSESSSKLEEDKDKKINEMS
LQQQVKQLQNGLAECKKQHQEVISVYRMHL LYAVQGQMDEDVQKVLKQILTMCKNQSQK K			l	ł	ł		KEVTKLKEALNSLSQLSYSTSSSKRQSQQLEA
1065		ł	1	İ			
1065			1				LYAVQGQMDEDVQKVLKQILTMCKNQSQK
APLPTGRAQMSPSGRLCLLTIVGLILPTRGOTL KDTTSSSSADATIMDIQVPTRAPDAYVTELQP TSPTPTWPADEITQPQTQQLGETDGPLVT DPETHKSTKAAHPTDDITTLSERPSSTDVQT DPQTLKPSGFHEDDPFTYDEHTLRKRGLLVA AVLFITGIILTSGKCRQLSRLCRNHCR AVLFITGIILTSGKCRQLSRLCRNHCR AVLFITGIILTSGKCRQLSRLCRNHCR RRRRRGVYSRKKMSLKSERRGHTVQSDLL CKKGCGYYGNPAWQGFCSKCWREYHKAR QRQQEDWELAERLQREEEAPASSQSSQGA QSLTFSKFEEKKTNEKTRVTTVKKFFSASSR VGSKKEIQEAKAPSPSINRGTEDRVSKEFIE FLKITHKTGQETYKQTKLFLEGMHYKRDLSIE EQSECAQDFYHNVAERMGTRGKVPPERVEKI MDQIEKYIMTRLYKYVFCPETTDDEKSKENE EQSECAQDFYHNVAERMGTRGKVPPERVEKI QRRRALRWYPQMLCVPYEDIPEVSDMVV KAITDIIEMDSKRVPRDKLACITKCSKHIFNAI KITINNEPSADDPLTTLIVTVLKGNPPRLQSNI QVITRFCNPSRLMTGEGGYPYCPEVSDMVV KAITDIIEMDSKRVPRDKLACITKCSKHIFNAI KITINNEPSAADDPLTLIVTVLKGNPPRLQSNI QVITRFCNPSRLMTGEGGYPYCPVEDIPEVSDMVV KAITDIIEMDSKRVPRDKLACITKCSKHIFNAI KITINNEPSAADDPLTPLIVTVLKGNPPRLQSNI QVITRFCNPSRLMTGEGGYPYDEIDPSVDMVV KAITDIIEMDSKRVPRDKLACITKCSKHIFNAI KITINNEPSAADDPLTPLIVTVLKGNPPRLQSNI QVITRFCNPSTMTGGGGYPYDQTVAG WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYKNLDLISQLNERQEBS WSPACLGVKQMYYQMYCHISGNEQBS WSPACLGVKQMYYQMYCHISGNEQBS WSPACLGVKQMYYQMYCHISGNEQBS WSPACLGVKQMYYQMYCHISGNEQBS WSPACLGVKQMYYQMYCHISGNEQBS WSPACLGVKQMYYQMYYQMYYQMYYQMYYQMYYQMYYQMYYQMYYQMY				[
KDTTSSSSADATIMDIQVPIRAPDAVYTELOP	1065	2415	Α	8841	3	663	AAATAASLSPRGCRLRTPSSDVGPSRAPPPSA
KDTTSSSSADATIMDIQVPIRAPDAVYTELOP		}		·			
TSPIPTWPADETPOPOTOTOQLEGTDOPLYT DPETHKSTKAAHPTDOTTITSPRSPSTDVQT DPETHKSTKAAHPTDOTTITSPRSPSTDVQT DPETHKSTKAAHPTDOTTITSPRSPSTDVQT DPETHKSTKAAHPTDOTTITSPRSPSTDVQT DPETHKSTKAAHPTDOTTITSPRSPSTDVQT DPETHKSTKAAHPTDOTTITSPRSPSTDVQT DPETHKSTKAAHPTDOTTITSPRSPSTDVQT DPETHKSTKAAHPTDOTTITSPRSPSTDVQT DPETHKSTKAAHPTDOTTITSPRSPSTDVQT DPETHKSTKAAHPTDOTTITSPRSPSTDVQT DPETHKSTKAAHPTDOTTITSPRSPSTDV DEVELOPE DPETHKSTKAAHPTDOTTITSPRSPSTDV DEVELOPE DEVELOPE DPETHKSTKAKPTYTOKKEPTSPASSR QKQQEDWELAERLQREEEEAFASSQSSQGA QSLTFSKFEEKKTNEKTKKVTTVKKFFSASSR VGSKKEIQEAKAFPSSINRGTSDRVSKEPTE FLKTFHKTGQETYKQTKLFLEGMHYKRDLSIE EGSECAQDPYFINVAERMQTISKVPERVSKII MDQIEKYIMTRLYKYVFCPETTDDEKKDLAI QKRIRALRWYTPQHLCYPVREDIPEVSDMVV KAITDIEMDSKRVPRDKLACHTKCSKHIFNAI KITKNEPSAADDELPTLIYIVLKGNPPRLQSNI QYTRFCNPSRMTGGGGYPYRCGAES WSPDACLGVKQMYKNLDLLSQLNERQERIM NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI KLDAQSLNLSQEDPERYMSQQTSPRKQEEIM NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI KPPNQPLAAIDSENVENDKLPPPLQPQVYAG SMREVGCGWLPVPIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKYHCAL LRQAWATKQDPISKKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISK LRQAWATKQDPISK LRQAWATKQDPISK LRQAWATKQDPISK LRQAWATKQD	j '		1				, ,
DPETHKSTKAAHPITODITTISERPSPSTOVOT				[ĺ	
DPOTILKPSGFHEIDPFFYDEHTIKRGILVA				1			
AVLITIGIIILTSGKCRQLSRLCRNHCR				ľ			
1066						l	
RRRRĞRVYSRKMSLKSERRGHYDQSDLL CKKGCGYYGNPAWQGFCSKCWREYHKAR QKQQQEDWELAERLQREEEEAFASSQSSQGA QSLTFSKFEEKKITIEKTITYTYKFFSASSS VGSKKEIQEAKAPSSSINRQTSIETDRVSKEFIE FLKTFHKTQQEIYKQTKLFLEGMHYKRDLSIE EQSECAQDFYHNVAERMQTRGKYPFERVEKI MDQIEKYMTRLYKYYCFPETTDDEKKDLAI QKRIRALRWYTPQMLCVPINEGKYPFERVEKI AUTOLICA QKRIRALRWYTPQMLCVPINEGKYPFERVEKI QKRIRALRWYTPQMLCVPINEGKYPFERVEKI QKRIRALRWYTPQMLCVPINEGUPIPEVSDMYV KATTDILEMDSKRYPPRDKLACTIKCSKHIPAJI KITKNEPASADDFLPTLIYIVLKGNPPRLQSNI QYTIRFCNPSRLMTGEDGYYFTINLCCAVAFIE KLDAQSINLSQEDFDRYMSGQTSPRKQBES WSPDACLGVKQMYKNLDLLSQLNERQERIM NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI KPPNQPLAAIDSENVENDKLPPPLQPQVYAG NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI KPPNQPLAAIDSENVENDKLPPPLQPQVYAG SNMREVGGOWLPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK KPPNQPLAAIDSENVENDKLPPPLQPQVYAG SNMREVGGOWLPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISK LRQAWATKYPTASTI LYDVQH LRQAWATKYPTASTI LYDVQH LRQAWATKYPTASTI LYDVQH LRQAWATKYPTAKKSSI LLRGKKKPPYRSGT QGCSVW GROPPLARCOGGNWILAQALICR GKQHILLKTEKSKLLSDISARWHTYRKKSSI LLRGKKKPPYRSGT QGCSVW GGGTOPSDAGWGCMLAQAMATCR GKQHLLKTEKSKLLSDISARWHTYRKKSSI GGGTOPSDAGWGCMLAQAMALCR GKQMLAQAMAMANAN LRGAWATCAGAMATCAGMMALAQALICR GKQHLLKTEKSKLLSDISARWHTYRKKSSI LRQAWATCAGAMAMANAN LRGAWATCAGAMAMANAN LRGAWATCAGAMAMANAN LRGAWATCAGAMAMANAN LRCAWATCAGAMAMANA	1066	2416	A	8853	3806	2204	
CKKGGGYYGNPAWQGFCSKCWREYHKAR QKQIQEDWELAERLQREEEAFASSQGA QSLTFSKFEEKKINEKTRKYTTVKKFFSASSGS QGQQEDWELAERLQREEEAFASSQSASQGA QSLTFSKFEEKKINEKTRKYTTVKKFFSASSR VGSKKEIQEAKAPSPSINRQTISIETDRVSKEFFF ELXTFHKTGQETYKGTIKLEGMHYKRDLSIE EQSECAQDFYHNVAERMQTRGKVPPERVEKI MDQIEKYIMRILYKYVTCPETTDDEKKDLAI QKRIRALRWYTPQMLCVPVNEDIPEVSDMVV KAITDIEMDSKRVPRDKLACITKCSKHIFNAI KITKNEPASADDFLPTLIYUTDDEKKDLAI KITKNEPASADDFLPTLIYUTDDEKKDLAI KITKNEPASADDFLPTLIYUTDDEKKDPPRLQSNI QYITRFCNPSRLMTGEDGYYFTNLCCAVAPIE KLDAQSLNLSQEDFDRYMSGGTSPRKQEAES WSPPACLGVKQMYKNDLSQLNERGERIM NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI KPPNQPLAAIDSENVENDKLPPPLQPQVYAG KPPNQPLAAIDSENVENDKLPPPLQPQVYAG SNMREVGGGWLPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LVLTQLSVTLSTVSVTPAVEVKGPVERDV LYLTQLSVTLSTVSVTPVAEVKGPVERDV EEGPPSYYDNQDPFATNWDDKSIRQAFIRKVF LVLTLQLSVTLSTVSVTFTVAEVKGFVRENV WTYYYSYAYFISLIVLSCCGDFRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG IITAVCFTVVIFSMGTRYDFTSCMGVLLVSM VVLFFEALICFFRNRIE,EIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVFAALNLYTDINI FLYLLTIGRAKE*PSSSSLCPLR WHGWPGPCP WHGSASCTSPLSCPQAPEKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVPPPVPAPQFOVEHPSPPPHPGVLPS GDMRSGGLPVLSPE GDMRSGGLPVLSPE GDMRSGGLPVLSPE GDMRSGGLPVLSPE GDMRSGGLPVLSPE LYLTHLPFLCKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRIHFMI*SIKLIMEETYLDIIKAVAYDRYTASII LNGEKLKVFPVRSGT'QGCSVWE LNGKKKVFPVRSGT'QGCSVWE GKQHLLKTEKSKLLSDISARWFTYRRKFSPI GGTOPSSDAGWGCMLAQAMLCRG GGTOPSSDAGWGCMLAQAMLCRG GGTOPSSDAGWGCMLAQCAMLACQAMLCRG GGTOPSSDAGWGCMLAQAMLCRG GGTOPSSDAGWGCMLAQCAMLACQAMLCRG GGTOPSSDAGWGCMLAQCAMLACQAMLCRG GGTOPSSDAGWGCMLAQAMLCRG GGTOPSSDAGWGCMLAQAMLCRG GGTOPSSDAGWGCMLAQAMLCRG GGTOPSSDAGWGCMLAQAMLCRG GGTOPSSDAGWGCMLAQAMLCRG GGTOPSSDAGWGCMLAQAMLCRG GGTOPSSDAGWGCMLAQAMLCRG GGTOPSSDAGWGCMLAQAMLCRG GGTOPSSDAGWGCMLAQAMLCRA GGTOPSSDAGWGCMLAQAMLCRA GGTOPSSDAGWGCMLAQAMLCRA GGT			l	"""	5000	-=-:	
QKQIQEDWELAERLQREEEEAFASSQSSQGA QSLTFSKFEEKKTNEKTRKYTKVKFFSASSR QSLTFSKFEEKKTNEKTRKYTVTVKKFFSASSR VGSKKEIQEAKAPSPSINRQTSIETDRVSKEFIE FLKTFIKTGQEIYKQTKLFLEGMHYKADLSIE EQSECAQDFYHNVAERMQTRGKVPPERVEKI MDQIEKYIMTRLYKYVFCPETTDDEKKDLAI QKRIRALRWYTPQMLCVPVNEDIPEVSDMV KAITDILEMDSKRVPRDKLACITKCSKHIPNAI KITKNEPASADDFLPTLIYIVLKGNPPRLQSNI QYTRPCNPSRLMTGEYYFTNLCCAVAFIE KLDAQSLNLSQEDFDRYMSQQTSPRKQEAES WSPDACLGVKQMYKNLDLLSQNERQERIM NEAKKLERDLDWTDGIAREVQDIVEKYPLEI KPPNQPLAAIDSENVENDKLPPPLQPQVYAG KPPNQPLAAIDSENVENDKLPPPLQPQVYAG SNMREVGCGWLVPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK PPPLQPQVAG SNMREVGCGWLVPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK LRQAWATKQDPISKKK PPPQGGYPQGPYPQGGYPQGP PPQGGYPQGPYPQGGYPQGP PPQGGYPQGPYPQGGYPQGP PPQGGYPQGPYPQGGYPQGP PPQGGYPQGPYPQGGYPQGP PPQGGYPQGPYPQGGYPQGP PPQGGYPQGPYPQGGYPQGP PPQGGYPQGPYPQGGYPQGP PPQGGYPQGPYPQGYPQGPYPQGPYPQGPYPQGPYPQGP	}		ľ	1			
OSLTESKFEEKTNEKTRKVTTVKKFFSASSR VGSKKEIQEAKAPSPSINQTSIETDRVSKEFIE FLYTFHKTGQEIYKQTKLFLEGMHYKRDLSIE EQSECAQDFYHNVAERMGTRGKVPPERVEKI MDQIEKYIMTRLYKYCEPETTDDEKKDLAI QKRIRALRWVTPQMLCVPVNEDIPEVSDMVV KAITDIIEMDSKRVPRDKLACITKCSKHIFNAI KITKNEPASADDFLITVIVLKGPPERVESDMVV KAITDIIEMDSKRVPRDKLACITKCSKHIFNAI KITKNEPASADDFLITVIVLKGPPRLQSNI QYITRFCNPSRLMTGEDGYYFTNLCCAVAFIE KLDAQSLNLSQEDFDRYMSQOTSPRKQEAES WSPDACLGVKQMYKNIDLLSQLNERQERIM NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI KPPNQPLAAIDSENVENDKLPPPLQPQVYAG SMREVGCGWLVPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISKK LRQAWATKQDPISK				1			
VGSKKEIQEAKAPSPSINRQTSIETDRVSKEFIE							
FLKTFHKTGQETYKQTKLFLEGMHYKRDLSIE							
EQSECAQDFYHNVÅERMQTRGKVPPERVEKI MDQIEKYIMTRLYKYVFCPETTDDEKKDLAI QKRIRALRWVTPQMLCVPVNEDIPEVSDMVV KAITDIEMDSKRVPRDKLACITKCSKHIFNAI KITKNEPASADDFLPTLIVIVLKGNPPRLQSNI QVITRFCNPSRLMTGEDGYYFINLCCAVAFIE KLDAQSLNLSQEDFDRYMSGQTSPRKQEAES WSPDACLGVKQMYKNLDLLSQLNERQERIM NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI KPPNQPLAAIDSENVENDKLPPPLQPQVYAG SNMREVGCGWLVPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK LVLTLQLSVTLSTVSVTFVAEVGFYRENV LVLTLQLSVTLSTVSVTFVAEVGFYRENV WTYYVSVAVFISLIVLSCCGDPRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG LVLTLQLSVTLSTVSVTFVAEVGFYRENV WTYYVSVAVFISLIVLSCCGDPRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG LVLTLQLSVTLSTVSVTFVAEVGFYRENV WTYYVSVAVFISLIVLSCCGDPRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG UTAVCTTVUIFSMQTRYDFTSCMGVLLVSM VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSSPESVYPAALNLYTDIINI LVLTLGLSTTLSTVSVTFVAEVGFYPPPPPGVLPS GDMRSGGLIPVLSPE GARGNTLYHLPRLCRKLNICRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPYRSGT*QGCSVWP LNGEKLKVFPYRSGT*QGCSVWP GKYLKKYEDQITIFTDYLEEYPDTDELVWIL GKYLKKYEDQITIFTDYLEEYPDTDELVWIL GKYLKKYEDQITIFTDYLEEYPDTDELVWIL GKYLKKYEDQITIFTDYLEEYPDTDELVWIL GKYLKKYEDGITIFTDYLEEYPDTDELVWIL GKYLKKYEDGITIFTDYLEEYPDTDELVWIL GKYLKKYEDGITIFTDYLEEYPDTDELVWIL GKYLKKYEDGITGPTYLEEYPDTDELVWIL GKYLKKYEDGITFTDYLEEYPDTDELVWIL GKYLKKYEDG			1				
MDQIEKYIMIRLYKYVFCPETTDDEKKDLAI QKRIRALRWVTPQMLCVPVNEDIPEVSDMVV							1 '
OKRIRALRWYTPQMLCVPVNEDIPEVSDMVV KAITDIIEMDSKRVPRDKLACITKCSKHIFNAI KITKNEPASADDFLPTLIYIVLKGNPPRLQSNI QYITRFCNPSRLMTGEDGYYFTNLCCAVAFIE KLDAQSLNLSQEDFDRYMSQGTSPRKQEAES WSPDACLGVKQMYKNLDLLSQLNERQERIM NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI KPPNQPLAAIDSENVENDKLPPPLQPQVYAG SNMREVGCGWLVPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK LRQAWATKQDPISKKK 1668 2418 A 8856 1530 1583 PCRPGMECNSMISVHCNL PVPQGYPQGPYPQGPYPQGPYPQGPYPQGPYPQGPYPQGP							
KAITDIEMDSKRVPRDKLACITKCSKHIFNAI KITKNEPASADDFLPTLIYIVLKGNPPRLQSNI QYITRFCNPSRLMTGEDGYYFTNLCCAVAFIE KLDAQSLNLSQEDFDR YMSGQTSPRKQEAES WSPDACLGVKQMYKNLDLLSQLNERQERIM NEAKKLEEKDLDWTDGIAREVQDIVEKYPLEI KPPNQPLAAIDSENVENDKLPPPLQPQVYAG EKPNQPLAAIDSENVENDKLPPPLQPQVYAG LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWEAEVGGSLEARS LRQAWATKQDFVIPAFWGGVPQGPQFPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQFYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQFYPQGGYPQFYPQGTYPYYVSYAYAYFISLIVLSCCGOFRRKHPWIL VALSVLTASLSYMVGMIASFYNTEAVIMAVG ITTAVCFTVVIPSMQTISTPTLEVIMAVG UTATVITGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHYFPPPPAPQPQEHPSPPPHPPGVLPS GDMRSGGLEPVLSPE GGTGPSSDAQWGCMLRCGGMMLAQALICRH GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH GGTGPSSDAGWGCMLRCGQMMLAQALICRH GGTGPSSDAGWGCMLRCGQMMLAQALICRH GGTGPSSDAGWGCMLRCGQMMLAQALICRH GGTGPSSDAGWGCMLRCGQMMLAQALICRH GGTGPSSDAGWGCMLRCGQMMLAQALICRH GGTGPSSDAGWGCMLRCGQMMLAQALICRH GGTGPSSDAGMGCMLRCGQMMLAQALICRH GGTGPSSDAGMGCMLRCGQMMLAQALICRH GGTGPSSDAGMGCMLRCGQMMLAQALICRH GGTGPSSDAGMGCMLRCGQMMLAQALICRH GGTGPSSDAGMGCMLRCGQMMLAQALICRH GGTGPSSDAGMGCMLRCGQMMLAQALICRH GGTGPSSDAGMGCMLRCGQMMLAQALICRH GGTGPSSDAGMGCMLRCGQMMLAQALICRH GGTGPSSDAGMGCMLRCGQMMLAQALICRH GGTGPSSDAGMGCMLRCGQMMLAQALICRH GATGPSSDAGMGCMLRCGQMMLAQALICRH GATGPSSDAGMGCMLRCGQMMALAQALICRH GATGPSSDAGMGCMACCGGMMLAQALICRH GATGPSSDAGMGCMAC			ļ	1			
KITKNEPASADDFLPTLIYIVLKGNPPRLQSNI QYITRFCNPSRLMTGEDGYYFTNLCAVAFIE KLDAQSLNLSQEDFDR YMSGQTSPRKQEAES WSPDACLGVKQMYKNLDLLSQLNERQERIM NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI EVPNQPLAAIDSENVENDKLPPPLQPQVYAG SMREVGGCWLVPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK 1068 2418		1	l	ŀ			VARDUE ACCEPTODOR ACCEPTOCATION A
QYITRFCNPSRLMTGEDGYYFTNLCCAVAFIE KLDAQSLNLSQEDFDR YMSQTSPRKQEAES WSPDACLGYKQMYKNLDLLSQLMERQERIM NEAKKLEKDLIDWTDGIAREVQDIVEKYPLEI KPPNQPLAAIDSENVENDKLPPPLQPQVYAG SNMREVGCGWLVPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK 1068 2418 A							KATI DILEMDSKKY PRDKLACI I KONDODI OGNI
RLDAQSLNLSQEDFDRYMSGQTSPRKQEAES WSPDACLGVKQMYKNLDLLSQLNERQERIM NEAKKLEKDLIWTDGIAREVQDIVEKYPLEI KPPNQPLAAIDSENVENDKLPPPLQPQVYAG KPPNQPLAAIDSENVENDKLPPPLQPQVYAG KPPNQPLAAIDSENVENDKLPPPLQPQVYAG KPPNQPLAAIDSENVENDKLPPPLQPQVYAG LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDFISKKK LRQAWATKQDPISKKA LRQAWATKQDPISKA LRQAWATKQDPISKA LRQAWATKQDPISKA LRQAWATKQDPISKA LRQAWATKQDPISKA LRQAWATKQDPISKA LRQAWATKQDPISKA LRQAWATKQDPISKA LRQAWATKQDPISKA LRQAWATKQDP							
WSPDACLGVKQMYKNLDLLSQLNERQERIM NEAKKLEKDLIDWTDGIAREVQDIVEKYPLE KPPNQPLAAIDSENVENDKLPPPLQPQVYAG SNMREVGCGWLVPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK SNMREVGCGWLVPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK LRQAWATKQDPISKKK SNMREVGCGWLVPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK LRQAWATKQDPISKKK SSSST 1530 1583 PCRPGMECNSMISVHCNL PCRPGMECNSMISVHCNL LRQAWATKQDPISKGPVPQGGYPQGPYPQGYPPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPYPQGGYPQGPPPQGPYPQGPYPQGPPPQGPYPQGPPPQGPYPQGPPPQGPYPQGPPPQGPYPQGPPPQGYPQGPPPQGPYPQGPPPQGPPPQGPPPQGPPPQGPPPQGPPPQGPPPQGPPPQGPPPQGPPQGPYPQGPPPQGPPPQGPPQG							
NEAKKLEKDLDWTDGIAREVQDIVEKYPLE KPPNQPLAAIDSENVENDKLPPLQPQVYAG 1067 2417		l	l				
1067 2417 A 8855 1372 1513 SNMREVGCGWLYPVIPAFWEAEVGGSLEARS LRQAWATKQDPISKKK 1068			ł				
1067		ļ])			l
LRQAWATKQDPISKKK	10/5	0415	 	0077	1070		
1068 2418	1067	241/	A	8855	1372	1513	1
1070							
1070 2420 A 8866 293 1675 PYPQGGYPQGPYPQGGYPQGPYPQGGYPQGP YPQSPFPPNPYGQPQFYPQGPYPQGGYPQGP YPQSPFPPNPYGQPQFPYPQGPYPQGYPQGP YPQSPFPNPYGQPQFPYPQGPYPQGYPQGP YPQSPFPNPYGQPQFPYPQGPYPQGYPQGP YPQSPFPNPYGQPQFPYPQGYPQGPYPQGP YPQSPFPNPYGQPQFPTFALKVF LVLTLQLSVTLSTVSVFTFVAEVKGFVERNV WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYILTIIGRAKF*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPGTTATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVPPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHFFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH							(
YPQSPFPPNPYGQPQVFPGQDPDSPQHGNYQ EEGPPSYYDNQDFPATNWDDKSIRQAFIRKVF LVLTLQLSVTLSTVSVFTTVAEVKGFVRENV WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFM*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHILKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH							
EEGPPSYYDNQDFPATNWDDKSIRQAFIRKVF LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV WTYYVSYAVFISLIVLSCCGDFRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFM*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP INGEKLKVFPVRSGT*QGCSVWP GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGGGGPSSDAGWGCMLRCGQMMLAQALICRH	1070	2420	Α	8866	293	1675	
LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SINKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWILL GKQHILLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH		Į.	1				
WTYYVSYAVFFISLIVLSCCGDFRRKHPWNL VALSVLTASLSYMVGMIASFYNTEAVIMAVG ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWILL GKQHILLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH		1	ļ] .			
VALSVLTASLSYMVGMIASFYNTEAVIMAVG ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM VVLFIFAILCIFRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWILL GKQHILLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH		1	1				LVLTLQLSVTLSTVSVFTFVAEVKGFVRENV
VALSVLTASLSYMVGMIASFYNTEAVIMAVG ITTAVCFTVVIFSMQTRYDFTSCMGVLLVSM VVLFIFAILCIFRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWILL GKQHILLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH		1	1				
VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSFPPHPPGVLPS GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHILKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH]]				VALSVLTASLSYMVGMIASFYNTEAVIMAVG
VVLFIFAILCIFIRNRILEIVYASLGALLFTCFLA VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSFPPHPPGVLPS GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHILKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH		}	1]	!
VDTQLLLGNKQLSLSPEEYVFAALNLYTDIINI FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGISQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFM*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH		-	ŀ				
FLYILTIIGRAKE*PSSSSLCPLRWHGWPGPCP WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HILLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSFPPHPPGVLPS GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SINKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHILKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH		1	1]			
WHGSASCTSPLSCPQAQPREKDASLQPSCMY TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHILKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH		(ĺ	1			
TADTSIWTRCGHSMAPLVLPPPPRGTKATFPC HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHILKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH				1			
HLLSTHCCMSPVCQPTPGTGGSTRSRGEGLSQ EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHILKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH		\	1				
EVRVHVFPPVPAPQPGVEHPSPPPHPPGVLPS GDMRSGGLIPVLSPE 1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHILLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH				i I			
GDMRSGGLIPVLSPE			ł				
1071 2421 A 8868 2 358 ARGNTLYHLPRLCRKLNLRWFSASTLYDVQH DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHILKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH							
DDKMGSNTFFKRNDCRYVMISCKADMAYDN VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH	1071	2421	Δ	8868	2	359	
VRHPFMI*SIKLIMEETYLNIIKAVYDRPTASII LNGEKLKVFPVRSGT*QGCSVWP 1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH	10/1	2421	^	0000	-	330	
1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH							
1072 2422 A 8870 33 658 MESVLSKYEDQITIFTDYLEEYPDTDELVWIL GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH			1	1		1	·
GKQHLLKTEKSKLLSDISARLWFTYRRKFSPI GGTGPSSDAGWGCMLRCGQMMLAQALICRH	1070	2422		0050		(50	
GGTGPSSDAGWGCMLRCGQMMLAQALICRH	1072	2422	I A	8870	53	028	
GGTGPSSDAGWGCMLRCGQMMLAQALICRH LGRDWSWEKQKEQPKEYQRILQCFLDRKDC			l				
		1	ĺ				GG I GPSSDAGWGCMLRCGQMMLAQALICRH
	L	J	<u> </u>	!	L	L	LGKDWSWEKQKEQPKEYQKILQCFLDRKDC

			Long	1 5 4: A 4	I Donathand and	Amino acid sequence (A=Alanine C=Cysteine,
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end nucleotide	D=Aspartic Acid, E=Glutamic Acid,
NO: of	NO: of	hod	ID NO:	beginning nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
nucl-	peptide	,	in USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
eotide	seq-		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
seq-	uence	i	914		acid residue	O=Glutamine, R=Arginine, S=Serine,
uence		ļ	914	ng to first amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	i			sequence	/=possible nucleotide deletion, \=possible
ł				peptide	1	nucleotide insertion
	ļ	ļ	 	sequence	 	CYSIHQMAQMGVGEGKSIGEWVLGPNTVAQ
					1	GV*KNLA/LFDEW\NSLGLVYVSM\DNPSGSIA
					1	RFPKKLCRVLPL\SADTAGLTGP
		.			410	DFSV*GDVDIEVTCPICLQLLTEPLSLNCGLRL
1073	2423	Α	8879	146	412	*OVCITA*IKESVIISGG*SSSPVCHTTFQPANL
	ļ	1				RTSRYLPT*SIKSLGPDEPQEG
					1,00	
1074	2424	Α	8884	67	435	HLQGRSIRTLQLTGENEKNCEVSERIRRSGPW
			1		1	KEISFGDYICHTFQGDCWADRSPLHEAAAHG
ĺ	1	1	ļ			RLLALKTLIAQGVNVNLWIL/DRVSSLHEACL
		İ.,				*GPVACAKPYWKMVPRHGGTVTGPPLLMV
1075	2425	Α	8896	1294	248	RSGDRNGLTHQLGGLSQGSRNQSYRSRSRSR
	1	1				SRERPSAPRGIPFASASSSVYYGSYSRPYGSDK
	}	1	}		Ì	PWPSLLDKEREESLRQKRLSERERIGELGAPE
ŀ		į .	ì	ĺ		VWGLSPKNPEPDSDEHTPVEDEEPKKSTTSAS
	1			İ		TSEEEKKKKSSRSKERSKKRRKKKSSKRKHK
						KYSEDSDSDSDSETDSSDEDNKRRAKKAKKK
		1]		}	EKKKKHRSKKYKKKRSKKSRKESSDSSSKES
İ		i	1		1	QEEFLENPWKDRTKAEEPSDLIGPEAPKTLTS
		1		1		QDDKPLNYGHALLPGEGAAMAEYVKAGKRI
l		1	1	ł		PRRGEIGLTR*RNCHHLNAQVM**VVSRHRR
				1	[MEAVRTAKREPESTVLMRREPLHPFNPRRET
		1	1		1	KERE
1076	2426	Α	8899	146	789	GRSTEAEKEPAFDERTGKGRRLPRAGEFHG*E
}		1	1 .		1	*APGPGPRSFQVSRKMPEE\PPGARKHPFSGKS
}	Ì	ì	1			FYLDLPAGKNLQFLTGAIQQLGGVIEGFLSKE
		1				VSYIVSSRREVKAESSGKSHRGCPSPSPSEVR
		1	1	1 '		VETSAMVDPKGSHPRPSRKPVDSVPLSRGKE
	1	1			İ	LLQKAIRNQK**CTVQQLSHCRLY\GEKTTAK
1	}	1		1 .		RSQREHVQQQSQEHGKWPDLKGPR
1077	2427	A	8901	352	3	AKIGAYKYIQELWRKKQSDVMHFLLRVRCW
			ł	1	Ĭ	QYPALHRAGTEWQLSALHRAPRSTQPDKAC
	i			1	1	RLGYKAKQGYIIYRICVRRGGWKCPVPKAVT
1					}	\YGKPVHHGVN*LKFAQSLQSVAEEQ
1078	2428	A	8905	536	781	ACPAENREVPEMAAGQAPHAGPGAGPGQPA
1	1	1				PALPFAATPGSRGQALCRGGRRRQHLHGPLH
1		ļ		1		RP*QAAPALHAGCQLAPHPPT
1079	2429	A	8912	121	376 ·	NLIWKLCVTERRLVILDNYDLASE/YEANKYI
1,		1				CNRIIQFKPGQDKYFTLGLPTGSTPL*CYPKLI
ł		1	ĺ			EYNKNGHLSFKYVKTFSMDEY
1080	2430	A	8920	381	1788	SSESPSDPGRMAMTWIVFSLWPLTVFMGHIG
1300	*****	^î.	3720	1	1	GHSLFSCEPITLRMCQDLPYNTTFMPNLLNHY
1	1	1			1	DQQTAALAMEPFHPMVNLDCSRDFRPFLCAL
		1			1	YAPICMEYGRVTLPCRRLCQRAYSECSKLME
1 -	1	1	1	}	1	MFGVPWPEDMECSRFPDCDEPYPRLVDLNLA
		1		1		GEPTEGAPVAVQRDYGFWCPRELKIDPDLGY
1	1		1	1		SFLHVRDCSPPCPNMYFRREELSFARYFIGLIS
i	1	i		1		IICLSATLFTFVTFLIDVTRFRYPERPIKCYAV
i]			1		WHMMVSLITTIGFLLEDRVACNA\SIPAQYKA
1	1	1		1		STVTQGSHNKACTMLFMILYFFTMAGSVWW
	1	1	}	1		VILTITWFLAAVPKWGSEAIEKKALLFHASA
1	1		1	1	Ī	WGIPGTLTIILLAMNKIEGDNISGVCFVGLYD
	1	1	1	1		VDALRYFVLAPLCLYVVVGVSLLLAGIISLNR
1	1		1	J		VRIEIPL*KENQDKLVKFMIRIGVFSILYLVPLL
	1	1		1	1	VVIGCYFYEQAYRGIWETTWIQERC
1000	1		10000	156	120	EERTKMSTGPDVKATVGDISSDGNLNVAQEE
1081	2431	A	8922	56	420	
1	1	1	1	1		CSRKGIVDEFFPLLSN*CIWTQPQGYPQSSYG
1			4			
1	1		İ	1		TLANFVF\CSVRHGLALILQLCNFSIYTQQMN
1082	2432	A	8923	355	1079	TLANFVFICSVRHGLALILQLCRFSIY I QQMN LSIAIPAMVNNTAPPSQPNASTERPST PFGTPSSTMAVVKNKCLMKGGKKGVKKKVV

SEQ ID NO: of NO: of nucleotide peptide sequence (SEQ ID NO: of nucleotide sequence (SEQ ID NO: of nucleotide sequence (SEQ ID NO: of nucleotide sequence (SEQ ID NO: of nucleotide sequence (SEQ ID NO: of nucleotide sequence (SEQ ID NO: of nucleotide sequence (SEQ ID NO: of nucleotide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotide location corresponding to last amino acid residue of peptide sequence (SEQ ID NO: of nucleotid	nic Acid, te, H=Histidine, Leucine, ine, P=Proline, S=Serine, =Tryptophan, *=Stop codon,
nucleotide sequence under the peptide under the peptide sequence under the peptide under the	ne, H=Histidine, Leucine, ine, P=Proline, S=Serine, =Tryptophan, *=Stop codon,
eotide sequence USSN 09/496 09/496 orresponding to last amino acid residue of peptide residue of peptide sequence USSN 09/496 orresponding to last amino acid residue of peptide sequence USSN 09/496 corresponding to last amino acid residue of peptide sequence T=Threonine, V=Valine, W= Y=Tyrosine, X=Unknown, 4 /=possible nucleotide deletion nucleotide insertion GPFSKKDQYDVKAPAMI	Leucine, ine, P=Proline, S=Serine, =Tryptophan, *=Stop codon,
sequence Sequence 09/496 Correspondi 10 10 10 10 10 10 10 1	ine, P=Proline, S=Serine, =Tryptophan, *=Stop codon,
uence 914 ng to first amino acid residue of peptide residue of peptide peptide sequence 914 ng to first amino acid residue of peptide residue of peptide sequence 914 ng to first amino acid residue of peptide residue of peptide y=Tyrosine, X=Unknown, 4 /=possible nucleotide deletion nucleotide insertion GPFSKKDQYDVKAPAMI	S=Serine, =Tryptophan, *=Stop codon,
amino acid residue of peptide	=Tryptophan, *=Stop codon,
residue of peptide /=possible nucleotide deletic sequence /=prossible nucleotide deletic nucleotide insertion /= GPFSKKDQYDVKAPAMI	=Stop codon,
peptide /=possible nucleotide deletic sequence nucleotide insertion GPFSKKDQYDVKAPAMI	
sequence nucleotide insertion GPFSKKDQYDVKAPAMI	on, rpossible
GPFSKKDQYDVKAPAMI	•
QGTQIASDGLKGLLFEVS	
FKLITEDVQDKNCLTNFY	YGMDLTCDKICSMV
EKWSTMIEAHVDVKTTD	
HNNQILKTSYA*HQQS/R	
DVFIRKVKMLENPGFER\	MELRGGGSSS
1083 2433 A 8948 28 385 LTWPQPHIPSCPAMSEET	LQSKLAAAKKKLP
WGAVQGSRAMSDLLLLI	LLDLTLLLLLMLLGF
AGYSGQLAGVAVSAGSP	
GWLLT/ESCSISPKLCSIA*	
1084 2434 A 8950 156 318 HYTPINTDTIENSENNKC	
WGGKRVQPFWKRVWQK	
1085 2435 A 8956 16 413 HMGQLGYFIQCWWECK	
*TIYTSYDTAIPIS/GI/YPK	
MFILAPFTATIKGKQLTCI	PLVEERIDYMWYS
HKYYIKVKRNL*VTITHN	IWVNLNILMFEIILW
YSHKYY	
1086 2436 A 8962 868 1026 H*KILQVGRAQRAHXSRI	
NPGARGCSEARLHRCTP/	
1087 2437 A 8985 58 330 LHVKHLGHFQLVFSEVIC	
*ERSVCAFHVCIQTYVCL	
FVYSVYGCGLCTCVCMD	
1088 2438 A 8989 394 404 N*KWILHVNVRIQSIFF/IK	(RNQK/INSHELKLD
KKFLDMMSNA*STKKHD	KLD/LIKFKT/LCSA
KYTVKRIKIHPTDLEKML	RNHLSDKD*YS/GV
YKDLSKLNRRKTE/S*/VK	KWVKDLSRYFIKE
VISMENKHKKIFSTS	
1089 2439 A 8991 60 329 MALTPESPSSFPGLAATG	SSVPEPPGGPNATI.
NSSWDSPTEPSSLEDLEA'	
GVEDNAYTLEVNSRYMR	
1090 2440 A 8996 2 351 SNITITLT*MKKYDNTFCV	
WQESKFIQAFWSKIQQYI	
GGYPGGTQSVFLTGVLVS	SCALAMWAY HALD
LLIAALFIIVQYWKQSKDI	
1091 2441 A 8997 97 456 YPLPVCSYLSGPRGEHWY	
LVSSRFKISKVIVVGDLSV	
AELGRVGPSLARWAGSR	
(
FHIFYVSVQNSISPSLSVSS	
RAAIIIHQHGQGPLGHGI	
1093 2443 A 9002 3 2745 ALLGLQQPAQSLILSRSSV	
TCPHICTVVNFKELAEHH	
AMCCLRYWYTPESWICG	
VKTFTAAGIKLIFFFDGM	
LKNNREISRIFHYIKSHKE	
VFTRFALKTLGQETLCSL	QEADYEVASYGLO
HNCLGILGEDTDYLIYDT	
TVMLCREKLCESLGLCVA	
PEGMFESFRYKCLSSYTS'	
VSDHISKVLYLYQGEKKL	
*RNGIISFTRT/INLHGFSKI	
RVOTPNPGKKFPCVOML	
GEKFPCIHLIPEPRQEVPTO	
PESRREVPMCSDPEPRQE	
MCTGPEARQEVPMCTDSI	
RQEVPMYTGSEPRQEVPM	
TGPESRQEVINTDPESRQ	
TOTESKŲEVLIKTOPESKŲ	YELLAIC I GERESKYE Y

NO. of cold body and control could be sequence of the cold of th	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	L Amino soid assures (A Alexies C C
							Amino acid sequence (A=Alanine C=Cysteine,
uence USSN Operation Ope			1.00	1	nucleotide		
1094 2444 A 9021 97 834 3902 1 337 1 1 1 1 1 1 1 1 1							
1094			1				
mino acid residue of peptide residue of peptide sequence peptide sequence peptide sequence per s			1	914			Q=Glutamine, R=Arginine, S=Serine.
residue of population popul			İ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
	j	ļ	J	ļ			Y=Tyrosine, X=Unknown, *=Stop codon.
PICTIPISKQEDSMCTHAERNQKLPVATIDEERK LEALMCINDEK(QEDTINOPERKQCUTNOPERK LEALMCINDEK(QEDTINOPERKQCUTNOPERK LEALMCINDEK(QEDTINOPERKQCUTNOPERK LEALMCINDEK(QEDTINOPERKQPT) CQDVTSTCLAVKEWPYYPONPLAHPILVRPL QMTIPGGTPSLKILM/JAQFEIQVRRLDTLLA CQDVTSTCLAVKEWPYYPONPLAHPILVRPL QMTIPGGTPSLKILM/JAQFEIQVRRLDTLLA CQDVTSTCLAVKEWPYYPONPLAHPILVRPL QMTIPGGTPSLKILM/JAQFEIQVRRLDTLLA CQDVTSTCLAVKEWPYYPONPLAHPILVRPL QMTIPGGTPSLKILM/JAQFEIQVRRLDTLLA CQDVTSTCLAVKEWPYPONPLAHPILVRPL QMTIPGGTPSLKILM/JAQFEIQVRRLDTLA CQDVTSTCLAVKEWPONPLANCILLITULVINSAGOL TUNNESTEDHWWNYPOGKLEHOKYLQGEKGYA VEVL/CRTK'ISARQIPQPGGSKQQLHEWPOCCCAVIVGAE THWAEPVSPLKHPVLAKKAITAIPDOLLEFV TEGSHPVEATKYPELBAITEDDLAVEMPSCCCEVIVGAE THWAEPVSPLKHPVLAKKAITAIPDOLLEFV TEGSHPVEATKYPELBAITEDDLAVEMPSCCCEVIVGAE THWAEPVSPLKHPVLAKKAITAIPDOLLEFV TEGSHPVEATKYPELBAITEDDLAVEMPSCCCEVIVGAE THWAEPVSPLKHPVLAKKAITAIPDOLLEFV TEGSHPVEATKYPELBAITEDDLAVEMPSCCCEVIVGAE THWAEPVSPLKHPVLAKKAITAIPDOLLEFV TEGSHPVEATKYPELBAITEDDLAVEMPSCCCEVIVGAE THWAEPVSPLKHPVLAKKAITAIPDOLLEFV TEGSHPVEATKYPELBAITEDDLAVEMPSCCCEVIVGAE THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV TEGSHPVEATKYPELBAITEDDLAVEMPSCCCEVIVGAE THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV TEGSHPVEATKYPELBAITEDDLAVEMPSCCCEVIVGAE THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV TEGSHPVEATKYPELBAITEDDLAVEMPSCCCEVIVGAE THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV TEGSHPVEATKYPELBAITEDDLAVEMPSCCCEVIVGAE THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV THEMAEPVSPLKHPVLAKKAITAIPDOLLEFV THEMAEPVSPLKHPVLAKTICSTST THEMAEPVSPLKHPVLAKTICSTST THEMAEPVSPLKHPVNCGGOTELLILLE THEMAEPVSPLKHPVLAKTICSTST THEMAEPVSPLKHPVNSTILLILLE THEMAEPVSPLKHPVNSTILLILLE THEMAEPVSPLKHPVNSTILLILLE THEMAEPVSPLKHPVNSTILLILLE THEMAEPVSPLKHPVNSTILLILLE THE			1	j	peptide	-	
LEALMCTNPEIRQEDPTNVCGEVKQQVTMVS			J		sequence		
DTELLKYARTHHYQAESYL-YNIMSGEEECS NTLEDELDQALPSQAPINSPIRGRYYSI.LIED CQDVTSTCLAVKSWFYYGRPRHPDL/RPI CQDVTSTCLAVKSWFYYGRPRHPDL/RPI CQDVTSTCLAVKSWFYYGRPRHPDL/RPI CQDVTSTCLAVKSWFYYGRPRHPDL/RPI CQDVTSTCLAVKSWFYYGRPRHPDL/RPI CQDVTSTCLAVKSWFYYGRPRHPDL/RPI CQDVTSTCLAVKSWFYYGRPRHPDL/RPI CQDVTSTCLAVKSWFYYGRPRHPDL/RPI CQDVTSTCLAVKSWFYYGRPRHPDL/RPI CQDVTSTCLAVKSWFYYGRGYGLOCLAUTHFYQ DYMPRAVQLOSLLTVALGLTLAVLNASACGF WKTSDFWFWNYSDGKHFQKYLQSEKGYA VEVLCRTK*ISAHQIPQFEGSRLQGLHEGEGT HWPSPLQLTPRREVGKTGLQL-PQDGLWV THWPSPLQLTPRREVGKTGLQL-PQDGLWV RWSPLCATKYRPICHURAVKATATHPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLLEFY THMAEPVSPLKHFVLAKKATTATPQLTSPLKHTATPATTATTATTATTATTATTATTATTATTATTATTATT							PICTOPISKQEDSMCTHAEINQKLPVATDFEFK
NTLEDELDQALPSQAFIYAPIRQRYYSILLED	}	ļ	ļ		ļ	ļ	LEALMCTNPEIKQEDPTNVGPEVKQQVTMVS
CQDVTSTCLAVKEWFVYPGPRENPLDITLA			l				DTEILKVARTHHVQAESYLVYNIMSSGEIECS
			l			ĺ	NTLEDELDQALPSQAFIYRPIRQRVYSLLLED
CFNLSSSRELQAVESPPQALCCLLIVILYONALON			l	1	ļ		CQDVTSTCLAVKEWFVYPGNPLRHPDLVRPL
DTI.CLEDLIJAFIAQALCLQGKSTSQLVALU_NSACGFP WKTSDFMPWAVPGGKLIVRGLTIL_VLNSACGFP WKTSDFMPWAVPGGKLIPIQKYLQSEKGYA WKTSDFMPWAVPGGKLIPIQKYLQSEKGYA WEVL/CRTK*ISAHQIPQFEGSRLQGLHEGEQT HHWPSPLGITFRREVGKTGLQLPQGLWV WEVL/CRTK*ISAHQIPQFEGSRLQGLHEGEQT HHWPSPLGITFRREVGKTGLQLPQGLWV WEVL/CRTK*ISAHQIPQFEGSRLQGLHEGEQT HHWPSPLGITFRREVGKTGLQLPQGLWV TEGSHFVEATYKNPELDRIATEDDL_VEMQGG THMAEPVSPLKHFVLAKKAITAFPQLLEFV TEGSHFVEATYKNPELDRIATEDDL_VEMQGG KDKLSIGEVLSRFMWVAFFGRTSSGKSSVI NAML.WDK VLPSGIGHITNCFLSVEGTODDK KDKLSIGEVLSRFMWVAFFGRTSSGKSSVI NAML.WDK VLPSGIGHITNCFLSVEGTODDK AGCLVRVFWFAKACALLRDDLVLVDGFGTD VTIELDSWIDKFCTKSSTREITNSGSDT VILNSRVEDFVPPEGAGREDFALRELAACW LLHRRARRSSALCPRPRSWGVSGGEGAGARE PHSSSCCLSAA/SHLSIGSPNMGARRIRRQ LAKEKIEGCHICTSVTFGEPQVFLGKDKAFTF DVYFDIDSQGEQHYQCEKLIEGCEGYANAT FAYGQTGAGKTYTMGTGFD TAYAGQTGAGKTYTMGTGTD TAYAGQTGAGKTYTMGTGFD TAYAGQTGAGKTYTMGTGTD TAYAGQTGAGKTYTMGTGAGAGKTYTMGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGA	ļ						QMTIPGGTPSLKILWLNQEPEIQVRRLDTLLA
DYNPRAYQLOSILVRGLTTL.VLYNSACGEYA WKTSDFMPWNYDFOGKLFIGHOXLOSEKGYA WEVL/CRTK*ISAHQIPQPEGSRLQQLHEGEQT HHWPSFLGTRREVOKTOLQLPQDGLWV							
WKTSDFM WNFPGKLFHOXTACRERGY	ì		ł				
VEVL/CRTK*ISAHQIPQPEGSSLQGLHEGGOT			ļ	1		}	DYINPKAVQLOSLLVKGLTTLVLVNSACGFP
1094							WELDLING WAS A HODODOOD OO HEDDOO
1094			1	ĺ			HUMBER OF TRABETORIES OF BODGE AND
Thmaepvsplkhfylakkaitaifdqllefy Tegshfylatkaitaifdqllefy Tegshfylatykhfylakkaitaifdqllefy Tegshfylatykhfylatiddlydemogy KDKLSHIGevlsrrhdrightsplothateddlydemogy KDKLSHIGevlsrrhdrightsplothateddlydemogy KDKLSHIGevlsrrhdrightsplothateddlydemogy Tythedswikskyktynqlahaihdbkdlk AgClyryfwrkakcallrddlyjdgfdd Tythedswikskyktynqlahaihdbkdlk AgClyryfwrkakcallrddlyjdgfdd Tythedswikstreitnisgdt Tythedswi	1094	2444	_	9021	07	824	
TEGSHFVEATYKNPELDRIATEDDIVEMOGY KDKLSHIGEVLSRRIMKVAFFORTSSGKSSVI NAMLWDK VLPSGIGHITNCFLSVEGTDGDKA VLMTEGSDEKSSVKTVNQLAHALMDKDLK AGCLVRVPWPKAKCALLARDDLVLVDGGTD VTTELDSWIDKFCTKSSTREITNSGSDT VTTELDSWIDKFCTKSSTREITNSGSDT VTTELDSWIDKFCTKSSTREITNSGSDT VTTELDSWIDKFCTKSSTREITNSGSDT LVLNSRVEDFVPFEGAGRTILFFALRPLACW LLHRRARRSSALCPRRSWOVSGGEGAGARE P*ITSSSCCLSAA/SHLSIQSPIMAGARRIRRYQ LAKEKIEGCHICTSVTPGEPQVFLGKDKAFTF DVVPDIDSQQEQIVIQCIEKLIEGCFEGYNATV FAYQQTGAGKTYTMGTGFD EDCLSPGGGNCESS*9CSVLRRILRQ EDCLSPGGGNCESS*9CSVLRRILRQ EDCLSPGGGNCESS*9CSVLRRILRQ EDCLSPGGGNCESS*9CSVLRRILRQ EDCLSPGGGNCESS*9CPYPAWITERDPV EDCLSPGGGNCESS*9CPYPAWITERDPV GKFFLVFLVETGFQCG*POLDLLTSRSACLG LPKCWDYRRFPAASIIFGTTFFINSK EVVVMSCEDDINISGSFYRNILKYLAFLCKATS TNPSQGPYHLWVPSHIFWQTTCGRLPHKTKQ G*AALDHLKVPDRFLPYDKKKQMAVSATLE VVRRFPRASIIFGTTFFINSK EVVVRKPPRKPAYLGHWAVSATLE VVRRFPRKPAYLGHWAVDWKYQMATA TMGEKRKVYYQKICYQKK UVRKPFRKPAYLGHWAVDWKYQMATA TMGEKRKVYYQKICYQKK UVRKPFRKPAYLGHWAVDWKYQMATA TMGEKRKVYYQKICYQKK UVRKPFRAYLGHWAVDWKYQMATA TMGEKRKVYYQKICYQKK UVRKPFRAYLGHWAVDWKYQMATA TMGEKRKVYYQKICYQKK UVRKFPRAYLGHWAVSATLE VVRRFPRASILECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHANRLYFK*RRGF ELFLERESLPVARVECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHANRLYFK*RRGF ELFLERESLPVARVECSGTISAHCNLC.PQSS DSPASAS*VAGITDMCRYTQLLIFHAS KTSMFWKFRDLHSSSHDTULLSFTVALMS UVEEPQQMDVLNWFKVQ RNL*HST*NVMDISKYVNILMEIN SKVLISLISKKPEQIVDFLKKKHDFVDLIKHIG TSAIMDLLLRLLTCTEPPQPRQDVLNWFKVQ RNL*HST*NVMDISKYVNILMEOLNISHSLL* LLLQCVLQWLNEEKIQRLVEIVHPSQEEDVS SLTSAIMDLLLRLLTCTEPPQPRQDVLNWFKVQ RNL*HST*NVMDISKYVNILMEOLNISHSLL* LLLQCVLQWLNEEKIQRLVEIVHPSQEEDVS SLY UVLQWYLNEKKIQRLVEIVHPSQEEDVS SLY UVLQWLNEKKIQRLVEIVHPSQEEDVS SLY UVLQWYLNEKKIQRLVEIVHPSQEEDVS SLAG AGMQUHLCWWCVNVKKFWGCGA AGMQUHLCWWCVNVKKFWEMS*YYLLKLSI AGMQUHLCWWCVNVKKFWEMS*YYLLKLSI AGMQUHLCWWCVNVKKFWEMS*YYLLKLSI AGMQUHLCWWCVNVKKFWEMS*YYLLKLSI AGMQUHLCWWCVNVKKFWEMS*YYLLKLSI AGMQUHLCWWCVNVKKFWEMS*YYLLKLSI AGMQUHLCWWCVNVKKFWEMS*YYLLK	1074	2444	1 ^	3021		034	
							TEGGHEVE ATVINDED DELATEDDE VEMOCY
NAML WDK.VLPSGIGHTINCELSVEGTOGDK.A		İ					
1095	l		ł				
AGCLVRVFWFKAKCALLRDDLVL/VDGFGTD							
1095							AGCLVRVFWPKAKCALLRDDLVLVDGPGTD
1095				.			VITELDSWIDKFCTKSSTREITNSGSDT
LIHRRARRSSALCPRPRSWQYSGGGAGARE	1095	2445	Α	9022	1	537	
P+ITSSSCCLSAA/SHLSIQSPMAGARRIRPQ							1
LAKEKIEGCHICTSVTPGEPQVFLGKDKAFTF							
DYVPDIDSQQEQIVIQCIEKLIEGCFEGYNATY							LAKEKIEGCHICTSVTPGEPQVFLGKDKAFTF
1096							DYVFDIDSQQEQIYIQCIEKLIEGCFEGYNATV
1097 2447 A 9032 716 357 ARSTGFWGEILWCGFLKRSLALSPRVKCSGAI LAHCNFRHAGFPLSCLSLPNRWEYRPPARP GKFFLVFLVETGFQC/G*DGLDLLTSRSACLG LPKCWDYRREPAASHIGVTFFINSK LAHCNFRHAGFPLSCLSLPNRWEYRPFARP GKFFLVFLVETGFQC/G*DGLDLLTSRSACLG LPKCWDYRREPAASHIGVTFFINSK LYLAFLCKRTS TNPSQGPYHLWYPSHIFWQTTCGRLPHKTKQ G*AALDHLKVFDRIPLPYDKKKQMAVSATLE VVRRKP*RKFAYLGHWAQKVDWKYQAMTA TMGEKRKVYQKICYQKK LYLAFLCKRTS TNPSQGPYHLWYPSHIFWQTTCGRLPHKTKQ G*AALDHLKVFDRIPLPYDKKKQMAVSATLE VVRRKP*RKFAYLGHWAQKVDWKYQAMTA TMGEKRKVYYQKICYQKK LYLHLPHNSVIAFLGISP RKIKTCPQNSCTSMLINAHHNQKWKKINI HSLJWK/TV*QFLKRLYLHILPHNSVIAFLGISP RKIKTCPQNSCTSMLINAHHNQKWKKINI LYSLAVENDE LESGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RGF LFFLEKVSNQFLSPSLLPVNFQGFVFAFLLLL FLLFEMESLPVARVECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RGF LFFLERKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLLFEMESLPVARVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLLLFHAS LYSLAVENDE LIFLAR LYSLAVENDE LIFLAR LYSLAVENDE LIFLAR LYSLAVENDE LIFLAR LYSLAVENDE LIKACLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLNWFKVQ RNL**HST*NVMDISKYVNLHWGLNKSHSLL** LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMVEIKSSPRFPIIKMTDITKCW*GCGGA AGMQUH/CW\WCVNVGKFWEMS*YYLLKSI							FAYGQT\GAGKTYTMGTGFD
EDCLSPGGGNCRES*SCPYTPAWITERDPV	1096	2446	Α	9029	1	285	FFFFNVCKSPKVPKPGCKEESTGTLFKNTLISL
1097							GQHSETPSLKKK\LAGYSGMCL*SQVLRRLRQ
LAHCNFRHAGFPPLSCLSLPNRWEYRRPARP	4400						EDCLSPGGGNCRES*SCPYTPAWITERDPV
GKFFLVFLVETGFQC/G*DGLDLLTSRSACLG	1097	2447	Α	9032	716	357	
LPKCWDYRREPAASIIFQTTFFINSK							
1098							
TNPSQGPYHLWVPSHIFWQTTCGRLPHKTKQ G*AALDHLKVFDRIPLPYDKKKQMAVSATLE VVRPKP*RKFAYLGHWAQKVDWKYQAMTA TMGEKRKVYYQKICYQKK 1099 2449 A 9043 185 372 IIFYSHQQCMRV/WQGCGDIETLIHCW*E*KII HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP RKIKTCPQNSCTSMLINAIHNDQKWKKINI 1100 2450 A 9045 763 584 RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCRLCPLSS DSFASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLTCIEPPQPRQDVLNWFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIIKMTDITKCW*GCGGA AGMQUH/CWWCVNVGKFWEMS*YYLLKLSI	1000	2440		0020	220	(70	
G*AALDHLKVFDRIPLPYDKKKQMAVSATLE VVRPKP*RKFAYLGHWAQKVDWKYQAMTA TMGEKRKVYYQKICYQKK 1099 2449 A 9043 185 372 IIFYSHQQCMRVYWQCGDIETLIHCW*E*KII HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP RKIKTCPQNSCTSMLINAIHNDQKWKKINI 1100 2450 A 9045 763 584 RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RRGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEHKSPRFPIKMTDITKCW*GCGGA AGMQU/H/CWWCVNVGKFWEMS*YYLLKLSI	1090	2446	А	9036	230	652	
VVRPKP*RKFAYLGHWAQKVDWKYQAMTA							
TMGEKRKVYYQKICYQKK					ľ		
1099 2449 A 9043 185 372 IIFYSHQQCMRV/WQGCGDIETLIHCW*E*KII HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP RKIKTCPQNSCTSMLINAIHNDQKWKKINI 1100 2450 A 9045 763 584 RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS DSPASAS*VAGITDMCRYTQLILFHAS IV*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HT*NVMDISKYVNLHWGLNKSHSLL*LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIKMTDITKCW*GCGA AGMQUH/CWWCVNVGKFWEMS*YYLLKLSI							TMGEKBKUKVOKICYOKK
HSL/WK/TV*QFLKRLYLHLPHNSVIAFLGISP RKIKTCPQNSCTSMLINAIHNDQKWKKINI 1100 2450 A 9045 763 584 RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RRGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIKMTDITKCW*GCGA AGMQUH/CWNWCVNVGKFWEMS*YYLLKLSI	1099	2449	A	9043	185	372	
RKIKTCPQNSCTSMLINAIHNDQKWKKINI 1100 2450 A 9045 763 584 RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPPHPANFLYFK*RRGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLNVWFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIKMTDITKCW*GCGA AGMQUH/CWWCVNVGKFWEMS*YYLLKLSI	/	/	•	7043	105	3/2	
1100 2450 A 9045 763 584 RQSLALSPRLECSGTISAHCRLCPLVFTPLSCL SLTSSWDYRRPPHPANFLYFK*RRGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLNVWFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW*GCGA AGMQUH/CW\WCVNVGKFWEMS*YYLLKLSI							
SLTSSWDYRRPPPHPANFLYFK*RRGF 1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTILEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLINDSFLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIKMTDITKCW*GCGGA AGMQUH/CWNWCVNVGKFWEMS*YYYLLKLSI	1100	2450	A	9045	763	584	ROSI AL SPRI ECSCTISANCEI CEI VETEI SCI
1101 2451 A 9050 275 2 LFFLRKVSNQFLSPSLLPVNFQGFVFAFLLLLL FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIJKSPRFPIKMTDITKCW*GCGGA AGMQUH/CW\WCVNVGKFWEMS*YYLLKLSI	,						SLTSSWDYRRPPPHPANFI VFK*RRGF
FLL/FEMESLPVA/RVECSGTISAHCNLCLPGSS DSPASAS*VAGITDMCRYTQLILFHAS DSPASAS*VAGITDMCRYTQLILFHAS DSPASAS*VAGITDMCRYTQLILFHAS SSPASAS*VAGITDMCRYTQLILFHAS DSPASAS*VAGITDMCRYTQLILFHAS DSPASAS*VAGITDMCRYTQLILFHAS SPASAS*VAGITDMCRYTQLILFHAS DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW*GCAGA AGMQUH/CWNWCVNVGKFWEMS*YYLLKLSI	1101	2451	A	9050	275	2	LFFLRKVSNOFLSPSLLPVNFOGEVFAFITITT
DSPASAS*VAGITDMCRYTQLILFHAS 1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLNWFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW*GCAGA AGMQUH/CWNWCVNVGKFWEMS*YYLLKLSI						-	
1102 2452 A 9053 449 1224 KTSMFWKFDLHSSSHIDTLLEREDVTLKELM DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLNWFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW*GCAGA AGMQUH/CWNWCVNVGKFWEMS*YYYLLKLSI			İ				
DEEDVLQECKAQNRKLIEFLLKAECLEDLVSF I*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLNVWFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIKMTDITKCW*GC*GA AGMQUH\CW\WCVNVGKFWEMS*YYLLKLSI	1102	2452	Ā	9053	449	1224	
IN*EEPPQDMDEKIRYKYPNISCELLTSDVSQM NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRILTCIEPPQPRQDVLNVWFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIKSPRFPIKMTDITKCW*GCGGA AGMQUH/CW\WCVNVGKFWEMS*YYYLLKLSI			- 1			ĺ	
NDRLGEDESLLMKLYSFLLNDSPLNPLLASFF SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPOPRQDVLNVWFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIKMTDITKCW*GCGGA AGMQUH/CW\WCVNVGKFWEMS*YYYLLKLSI							
SKVLSILISRKPEQIVDFLKKKHDFVDLIIKHIG TSAIMDLLLRLLTCIEPPQPRQDVLNVWFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIKMTDITKCW*GCGGA AGMQUH/CW\WCVNVGKFWEMS*YYLLKLSI			ļ				
TSAIMDLLLRLTCIEPPQPRQDVLN/WFKVQ RNL*HST*NVMDISKYVNLHWGLNKSHSLL* LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW*GC\GA AGMQUH/CW\WCVNVGKFWEMS*YYLLKLSI			ſ	ļ	ĺ	ſ	
RNL+HST+NVMDISKYVNLHWGLNKSHSLL+ LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW+GC\GA AGMQUH\CW\WCVNVGKFWEMS*YYLLKLSI				ļ	1		•
LLLQCVLQWLNEEKIIQRLVEIVHPSQEEDVS SLV 1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW+GC\GA AGMQUH/CW\WCVNVGKFWEMS*YYLLKLSI				1	j		
1103 2453 A 9058 403 3 GLHVYDFQVYREHILTLNVKKCSVSFWGLRE WLYLQMYEIIKSPRFPIIKMTDITKCW+GC\GA AGMQUH\CW\WCVNVGKFWEMS*YYLLKLSI				1		l	
WLYLQMYEIIKSPRFPIKMTDITKCW+GC\GA AGMQVH/CW\WCVNVGKFWEMS+YYLLKLSI							
WLYLQMYEIIKSPRFPIIKMTDITKCW+GCAGA AGMQI/H/CW\WCVNVGKFWEMS+YYLLKLSI	1103	2453	A	9058	403	3	GLHVYDFQVYREHILTLNVKKCSVSFWGLRE
	- 1		- 1	ļ	J	J	WLYLQMYEIIKSPRFPIIKMTDITKCW*GC\GA
	Ì	1		ļ	· 1		
							ST/PYDPAIPLLGIYL*ETRVYIHPKTCMRMLIA

COTO ID	L OF O III	T 3.7-4	1.000	T N	T	
SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	nou	in in	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid,
eotide	seq-		USSN	location	corresponding	F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine.
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		ļ	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		ļ	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	1	i	1	peptide	1	/=possible nucleotide deletion, \=possible
	1			sequence		nucleotide insertion
	Ī	1				APFVLAVNC
1104	2454	Α	9064	75	393	KWLFSSLNITGRGDIIGHLKWLDCR\NCSSFPI
1						KRNRQTHSTESNKLKAGHSFGYN*LIH*NS\V
ł			İ			KTDCGCGANSKGVVVVMKV\KTAQQKQTTS
		<u> </u>	<u> </u>			YMQIGTTKNSRAT
1105	2455	Α	9065	366	778	DLLILRNLAFPELKRRNCISRFYLAYHLHKIYS
1			1			RSILLCNNCSGFYILSL*QYDVFFFNYFFFRDR
				<u> </u>		AWPCCPGWSAAWLTIVILAHYRRPGLERSCC
ł	ļ					LSLSSSWDHRRVPPCPANF*/YFSMGFTAFPRL
1104	3456	 	0000			VLNS*TQGI
1106	2456	Α	9083	673	816	ESGSLIH*WWENKPAQPLWWEI*QHVQKLPT
1107	0450	<u> </u>	0006			HFPCDPAIPLLGICPED
1107	2457	A	9086	580	18	KPSSGSFIRAIYIFLSTAHVPALFSVLVRTKLT*
		1				AFSQSSVLWAHKQQKTSLSLVIR/ERLQIKTA
		1				VRENFLPIRLAKILKLDNVKCWQG/SGSNMSL
	İ	1				I/HCWWEYNVIHIIWNSVTFPRKVEHVYITYA
į		ļ				PEISVR*IHGGLPTLVHQETHTSVFRGAPSVIP ETR\CRPTKESINKLLHIYTMEHYGDENK
1108	2458	Α	9093	540	1	GGNDCSVTPTTEPGRKEIT*KRKF*EKTDRLP
	2.50	1.	7075	540	'	GA/PPSRTPPTPYPCPHGDRLLPPSRPLPAGPA
						SAFPPAERSRGHRRASL*RARWSAAVPRRSA
		l			ľ	GSASEPVQSRWLRLPVGSDSPPAVPVRVCPAP
İ						DSRPAAPGSRLPDPGLDSPAPSRTPSSSVD+GG
			1			QRPPPPSGDSLSPPGCCRY
1109	2459	A	9099	1255	1425	HESYHVNPNLCNPVAPTSGAHSIG*KWPSWL
	1	l				GAVAHSCNPSTLVGRGGRITRGQELR
1110	2460	Α	9103	242	70	EEQFFFFAVGMFP*VDFLAPASGELWDRLRLT
]				•		CSRPFTRHQSFGLAFLRVCSSLDSLDDSVVGP
						SALLSSVL/NQGGRNVLEAREAAKHPTI*RQS
						LLRKQRNKRMAIP
1111	2461	Α	9110	189	121	SFLSVRLECNGAIMAHCALPLPG
1112	2462	A	9113	100	910	RRRGGGSRPRRTPVPAPGPGPSFGMDVRFYP
						AAAGDPASLDFAQCLGYYGYSKFGNNNNYM
						NMAEANNAFFAASEQTFHTPSLGDEEFEIPPIT
			1 1			PPPESDPALGMPDVLLPFQALSDPLPSQGSEFT
						PQFPPQSLDLPSITISRNLVEQDGVLHSSGLHM
						DQSHTQVSQYRQDPSLIMR\PSST*PDAARSG
						VMPPAQLTTINQSQLSAQLGLNLGGASMPHT SPSPPASKSATPSPSSSINEEDADEANRAIGEK
		•				RAAPDSGKKPKTPKK
1113	2463	A	9120	3452	3051	FLRPSFALVPQAGVQWCALSWLQPPSPRFK*F
					-551	SCLSLPSSWDYRHVPPRPANFFVLLVETGFLH
						VGQAGHEPLTSGDPPASASQSAGITGVSHQA
ľ	•		ľ - 1		• • • • •	WPSFFIFSRDTVLLCCSGWSRTSGLKQSACLS
						LLKCWDY
1114	2464	Α	9122	152	377	NQLPLQQWTFFIYETGFCSVAQAGVQCRDHS
						SLHP*PPG\SSDPPAPPS*VLGITGQRYHACLII
				i	4	YLYVQTVPQRV
1115	2465	Α	9124	553	981	QRPLLRQQLGSWPTCRSLEGDLASPW**RLPG
				ļ	}	SPRMRRSGT/ATLNLPLSPQGTVRTAVEFQVM
				ļ	}	TQTQSLSFLLGSSASLDCGFSMAPGLDLISVE
					ł	WRLQHKGRGRGDLHLPDHHLSVPSSADHPA
					1	QQPSQFNGRNLYFLPLFR
1116	2466	Α	9135	48 .	410	SASHEPAEHDGGADSLSASQPPRPAGRPAGA
				1	- 1	QHVHVPPWTDVLAGQDRRAPTAGDGAPWP
				l		APGGHVPSTRPHDPAEFHADEAAGRGGRGLQ
					ļ	PAAPHALPAGLPHGPPAPA/PAEGGGTP*GSA
				1	Į	GAGGP*GSPAGRACGAAGCRPRPPRPAASSA
						*NSAGS*GLVEGT*PPGAGHGAPSPAVGARLS
			_			

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location corresponding to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion CPARTSVQGGTWTC*APAGRPAGLGGWEAE RESAPPSCSAGS*DAD*GAEPWGAGSRSWGS KSGHWAKECLQPRIPPRPCPICVGPHWKSDCP TCPGAVPRAPGTLPQGSLTDSFPDLLSLVAED *CCLMASEASWTINELWVTLTVEGKSVP/CL NTEATHSTLPSFQGPVSLASITVVGIDGQASKP LKTPQLWCQLGQYSFMHYFLVIPTCPVPLLG* GILTKLSAFUTIPRLQPHLIAALSPSS
1118	2468	A	9154	471	2	AAGQVVVEVTSHLYLCITSDAAGLRLLPPAES ERGEGGHCPAEAPLPPRPQYCLAKHPLLRKLP EEKIKLDPYLTQHTKINSKQIKYLS/VRAKTTQ LVEGNIGVNLQNTELKQH*INGFLDTTPEAQE TKEKTNKLNFIKKVKRQLAEWEKIFQIA
1119	2470	A	9155	124	3187	ACPRLARRRRRVRSLRRRRGWLRARWSRGQ NNMAARRITQETFDAVLQEKAKRYHMDASG EAVSETLOFKAQDLLRAVPRSRAEMYDDVHS DGRYSLSGVAHSRDAGRESLRSDVFSGPSFR SSNPSISDDSYFRKECGRDLEFSHSNSRDQVIG HRKLGHFRSQDWKFALRGSWEQDFGHPVSQ ESSWSQEYSFGPSAVLGDFGSSRLIEKECLEK ESRDYDVDHPGEADSV/LRGGSQVQARGRAL NIVDQEGSLLGKGETQGLLTAKGGVGKLVTL RNVSTKKIPTVNRITPKTQGTNQIQKNTPSPD VTLGTNPGTEDIQFPIQKIPLGLDLKNLRLPRR KMSFDIIDKSDVFSRFGIEIKWAGFHTIKDDIK FSQLFQTLFELETETCAKMLASFKCSLKPEHR DFCFFTIKFLKHSALKTPRVDNEFLNMLLDKG AVKTKNCFFEIIKPFDKYIMRLQDRLLKSVTP LLMACNAYELSVKMKTLSNPLDLALALETTN SLCRKSLALLGQTFSLASSFRQEKIL*AVGLQ DIAPSPAAFPNFEDSTLFGREYIDHLKAWLVS SGCPLQVKKAEPEPMREEEKMIPPTKPEIQAK APSSLSDAVPQRADHRVVGTIDQLVKRVIEGS LSPKERTLLKEDPAYWFLSDENSLEYKYYKL KLAEMQRMSENLRGADQKPTSADCAVRAML YSRAVRNLKKKLLPWQRRGLLRAQGLRGI WKARRAITTGTQTLLFLRAPGLKHHGRQAPG LSQAKPSLPDRNDAAKDCPPDPVGPSPQDPSL EASGPSPKPAGVDISEAPQTSSPCPSADIDMKT METAEKLARFVAQVGPEIEQFSIENSTDNPDL WFLHDQNSSAFKFYRKKVFELCPSICFTSSPH NLHTGGGDTTGSQESPVDLMEGEAEFEDEPP PREAELESPEVMPEEEDDDEDGGEEAPAPG GAGKSEGSTPADGLYGARGHURACHURACH WKARRANTGGQESPVDLMEGEAEFEDEPP PREAELESPEVMPEEEDDDEDGGEEAPAPG GAGKSEGSTPADGLYGARGHURACHURACH WASRAVROLGE PKGECPPVGTVASSTVLGWWAVRVRRDRWR HFNPKEFCAPLQNVSRHSCFPVV
1121	2471	A	9166	272	523	PMSSLQGCFYTFKCIIFKGIFLLLISNLIAF**EK
		1				V/CSHITDSLKFIGKGWVGMVTHACNPGTLG G*GGWIA*VREFETSLGNM
1122	2472	С	9170	442	236	MNRRRFLRPADCHSGMRGTENGACSEGESQI HCGAGGEGVQLVHVVNQPENGCLQFDSTHIT FSKRQN*
1123	2473	A	9171	10	423	MVDRSPLLTSVIIFYLAIGAAIFEVLEEPHWKE AKKNYYTQKLHLLKEFPCLGQEGLDKILEVV SDAAGQGVAITGNQTFNNWNWPNAMIFAAT VITTIGYGNVASKTPGGRLFCGFYGLFGVPFC LTWINALGKFFG

	1 000 10	1.65	Tono-	T - 11 - 11 - 1		
SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	ļ	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496 914	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ŀ	914	ng to first amino acid	acid residue	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan.
]	1	l	1	residue of	of peptide sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
		1	Ì		sequence	
İ				peptide		/=possible nucleotide deletion, \=possible
1124	2474	A	9173	sequence	374	nucleotide insertion
1124	24/4	Ι Α	91/3	13	374	GPSPSLLVLLPQEPGGTGTPVRAGAGAGMWL
1	1		1			WEDQGGLLGPFSFLMLMLLLETRNPVNACLL
	ļ	1	ł		•	TGSLFVLLGVFSFEPVPSCRALQELKPRDRISA IAHRGGRHDPPENTLGAIR/QGS**WSNRR
1125	2475	A	9179	704	188	ESSGLLFQCFQGIHVQKLTLQARPTLFSWWL
1123	2473	^	7179	104	100	CSKPPKETGELENAESGGDGGRRGGKODNV
İ	1	ļ	1		l	AWWRRM\QKG\DFPWDDEDFPQSGPFGGQA LPMGFFYLYFRDPGREITWKHFVQYYLARGL
	İ	l				
ĺ		ĺ	Ì			VDRLEVVNKQSVRVIPAPGTSSEVRGEFKAE YCRHKFISCKNVVFYFFQ
1126	2476		9183	153	233	MEYMAESTDRSPGHILCCECGVPISPN
1127	2477	A	9185	1	321	LTGOLGSILLRVFSKSRAGLGARKLKAYRTM
1127	24//	Α .	9163	1 1	321	
		İ		1		EYMAESTDRSPGHILCCECGVPISPNPAQY\CV
				ì		ACLRSSFHIYHCIPKLFIHPFSKTSSSAFITPSHY LTFFSTIS
1128	2478	A	9186	183	847	VLKFLLLOTMDEOSOGMOGPPVPOFOPOKAL
11120	24/0	^	9100	163	047	RPDMGYNTLANFRIEKKIGRGO\FSEVYRAAC
		1			j	L\LDGVPVALKKVQIFDLMDAKARADCIKEID
ļ		j]			
				1		LLKQLNHPNVIKYYASFIEDNELNIVLELADA
			l	İ		GDLSRMIKHFKKQKRLIPERTVWKYFVQLCS ALEHMHSRRVMHRDIKPANVFITATGVVKLG
						DLGLGRFFSSKTTAAHSLVGTPYYMSPERIHD
1		1		ł		NG
1129	2479	A	9190	1	370	GTSWKIPSAAVSESSPNGAAYASGLPCGVRG
1129	2419	^	3130	. .	370	PPWAGLALLPSPTLMALLRRPTVSSDLDNIDT
						RATTIKIRVVATITRARIEDMRHSATALTRPD
				•		ATTAQIPKLPVTTVCNRRANPGIPPSVL
1130	2480	A -	9194	131	487	AYLKRLPVPESITGFARLTVSEWLRLLPFLGV
1130	2400	^	7174	1,31	707	LALLGYLAVRPFLPKKKQQKDSLINLKIQKEN
ļ			}]		PKVVNEINIEDLCLTKAAYCRCWRSKTFPAC
ļ		ļ				DGSHNKHNELTGDNVGPLILKKKE
1131	2481	A -	9201	184	605	KELVDEKSERGRAMDPVSQLASAGTFRVLKE
''''	2401	^	7201	104	005	PLAFLRALELLFAIFAFATCGGYSGGLRLSVD
ł	l	ł	1			CVNKTESNLSIDIAFAYPFRLHQVTFEG\PTCE
						GKERHKLALIGDSSSSAEFFGTVAGFAFLYSL
i		,				AATGVYIFFQNKY
1132	2482	A	9206	1	852	GGGRAGAGSRDMGSTDSKLNFRKAVIOLTTK
1	1	- -	1	Ī -		TQPVEATDDAFWDQFWADTATSVQDVFALV
1						PAAEIRAVREESPSNLATLCYKAVEKLVOGA
1	ļ	l		1		ESGCHSEKEKQIVLNCSRLLTRVLPYIFEDPD
1	1	l	}			WRGFFWSTVPGAGRGGGGEEDDEHARPLAE
	1					SLLLAIADLLFCPDFTVQSHRRSTVDSAEDVH
	ļ					SLDSCEYIWEAGVGFAHSPOPNYIHDMNRME
ŀ	1	ľ		· ·	'	LLKLLLTCFSEAMYLPPAPESWQH/RTHWFSS
1	İ	1	i			FVSSENRHALPLFTSLLNTVCAYDPVEYGIPY
	1		1			NHLY
1133	2483	A	9208	1165	1463	GPRARVQGFSGADIVKFMALGSMYLVLTLIV
]	J	1			AKVLRGAEPCCGPLKNRVLRPCPLP/VPLPPP
		1	1			HPQPSRGNPVGCLPTYKVVYKLLSWPLHSNS
						NVÝFIV
1134	2484	Α	9210	66	1586	MAGAGPKRRALSAPVAEEKEEAREKIMAAK
1	1	Ì	l		٠,	RADGAAPAGEGEGVTLQGNITLLKGVAVIVV
						AIMGSGIFVTPTGVLKEAGSPGLALVVWAAC
		1	l			GVFSIVGALCYAELGTTISKSGGDYAYMLDV
	1	1	1		·	YGSLPAFLKLWIELLIIRPSSQYIVALVFATYL
1	1	1	1			LKPLFPTCPVPEEAAKLVACLCVLLLTAVNC
	!		1			YSVKAATRVQDAFAAAKLLALALIILLGFVQI
l	1					GKGDVSNLDPNFSFEGTKLDVGNIVLALYSG
		L				LFAYGGWNYLNFVTEEMINPYRNLPLAIIISLP

CEO ID	SEQ ID	Med	Lego	Daniel and	15 37 1 1	
SEQ ID NO: of	NO: of	Met	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in in	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine.
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		}	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ì				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	•	/=possible nucleotide deletion, \=possible
		ĺ		sequence		nucleotide insertion
						IVTLVYVLTNLAYFTTLSTEQMLSSEAVAVDF
	1	ĺ			Í	GNYHLGVMSWIIPVFVGLSCFGSVNGSLFTSS
		İ				RLFFVGSREGHLPSILSMIHPQLLTPVPSLVFT
}					Ì	CVMTLFYAFSKDIFSVINFFSFFNWLCVALAII
		1	}		J	GMIWLRHRKPELERPIKVNLALPVFFILACLF
	1	1				LIAVSFWKTTPWSVASDFTIILSGLPVYFFGV
1126	2405		0016		1	WWKNKPKWAPPGHLSPRPSCVRSSCMVVPQ
1135	2485	A	9216	40	410	RDRLPPAYFCRPVVCVVTALDVG\SPESQEM
	ł					DLVAFEDVAVNFTQEEWSLLDPSQKNLYREV
		ļ				MQETLRNLASIGEKWKDQNIEDQYKNPRNNL
1136	2486	A	9223	3	002	RSLLGERVDENTEENHCGETSSQIPDDTLNK
1130	2460	A	9223	3	983	RRRRSRYRRCSRFPRPGPLAVSMPHAFKPG
]			DLVFAKMKGYPHWPARIDDIADGAVKPPPN
	ł		ŀ			KYPIFFFGTHETAFLGPKDLFPYDKCKDKYGK
l	į	ĺ	i :			PNKRKGFNEGLWEIQNNPHASYSAPPPVSSSD
	1		ļ			SEAPEANPADGSDADEDDEG\RGVMAVTAVT ATAASDRMESDSDSDKSSDNSGLKRKTPALK
						MSVSKRARKASSDLDQASVSPSEEENSESSSE
						SEKTSDQDFTPEKKAAVRAPRRGPLGGRKKK
			ł I			APSASDSDSKADSDGAKPEPVAMARSASSSSS
						SSSSSDSDVSVKKPPRGRKPAEKPLPKPRGRK
						PKPERPPSSSSSD
1137	2487	A	9229	21	239	LFPRLECRDPVTVNCTLNLPGSKNAPTTASOV
			1			GSTWNYRGGLPHPTNFFVKTGFRCSQAGLKL
	ļ.			,		RGSREPPAWA
1138	2488	A	9231	1664	2	TRSVGVNTCEVGVVTEPECLGPCEPGTSVNL
						EGIVWHETEEGVLVVNVTWRNKTYVGTLLD
						CTKHDWAPPRFCESPTSDLEMRGGRGRGKR
	1			·		ARSAAAAPGSEASFTESRGLQNKNRGGANGK
						GRRGSLNASGRRTPPNCAAEDIKASPSSTNKR
			1			KNKPPMELDLNSSSEDNKPGKRVRTNSRSTP
						TTPQGKPETTFLDQGCSSPVLIDCPHPNCNKK
		ì				YKHINGLRYHQAHAHLDPENKLEFEPDSEDK
						ISDCEEGLSNVALECSEPSTSVSAYDQLKAPA
		[ļ		SPGAGNPPGTPKGKRELMSNGPGSIIGAKAGK
						NSGKKKGLNNELNNLPVISNMTAALDSCSAA
						DGSLAAEMPKLEAEGLIDKKNLGDKEKGKK
		J	}			ANNCKTDKN\PSKLKSARPIAPAPAPTPPQLIA
			1	ĺ		IPTATFTTTTTGTIPGLPSLTTTVVQATPKSPPL
		1	İ	.		KPIQPKPTIMGEPITVNPALVSLKDKKKKEKR KLKDKEGKETGSPKMDAKLGKLEDSKGASK
		İ				DLPGHFLKDHLNKNEGLANGLSESOESRMAS
	<u>. </u>			. [IKAEADKVYTFTDNAPSPSIGS
1139	2489	A	9234	207	443	TRRGOPWRRAAAAGILPGREAAACLPSC/AS
						VTAAVSGLLVGYELGIISGALLQIKTLLALSC
		ł	l	ļ		HEQEMGVSSLVIGALL
1140	2490	A	9238	248	328	MAQGNNYGQTSNGVADESPNMLVYRKV
1141	2491	A	9242	2	535	FVEAAVKMLGSLVLRRKALAPRLLLRLLRSP
İ		- 1	ï Ì	1		TLRGHGGASGRNVTTGSLGEPQWLRVATGG
		į		1	i	RPGTSPALFSGRGAATGGRQGGRFDTKCLAA
				İ		ATWGRLPGPEETLPGQDSWNGVPSRAGLGM
		l		ļ		WPWAAALVVHCYSKSPSNKDAALLEAARAO
		ł	}	ì	ļ	\NMQEVSRNRCALLHSAAVQEYGYGN
1142	2492	A	9245	157	466	HLCFWFFVGLFLPEQQIMLFATLLRMAQGCD
			E .		1	
		1	1	l	Į.	FALGNUFLNII I KAQAVI KEKLUKLUFIKIK IC. I
			!			FALGNDFLNITTKAQA/TKEKLDKLDFIKIKTC CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF
1143	2493	A	9247	264	115	CTSMDAIEKTEPLTKWTKAFVSHVSYKRLLF

SEQ ID	SEQ ID	Met	SEQ	Predicted	I Deading and	I Amino cold comment (A. Albeiro O. C
NO: of	NO: of	hod	ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide	1104	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
•		1		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	1	l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
	ł	İ	1	peptide	ł ⁻	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1144	2494	A	9260	1	401	KKVPGRLSEMSFSLNFTLPANTTSSPVT\DCGP
	1	,	ļ	ļ	}	SLGLAAGIPLLVATALLVALLFTLIHRRRSSIE
	Ì	1		ĺ		AMEESDRPCEISEIDDNPKISENPRRSPTHEKN
	İ	1	ł		ł	TMGAQEAHIYVKTVAGSEEPVHDRYRPTIEM
1145	2406	<u> </u>	0074		 	ERRR
1145	2495	Α	9264	175	411	METIWIYQFRLIEIGDSTVGKSCLLHRFTQGRF
	Ì			ŀ		PGLRSPACDPTVGVDFFSRLLEIEPGKRIKLLL
1146	2496	A	9277	592	014	WDTAGQERFISIT
1140	2490	A	92//	392	814	MFTYLEGREGIKSQPKMEPHSVT\RLECSGMI
		ļ				SAHCSLNLPGTSDSPASASR/VAGTTGMRHHA
1147	2497	A	9279	1255	2	WLIFAFLVETGF FRRGRRGEEEKEEEEEEEGWVNGMENSHPP
114/	2471	^	3213	1233	4	HHHHOOPPPOPGPSGERRNHHWRSYKLMIDP
		İ				
ļ						ALKKGHHKLYRYDGQHFSLAMSSNRPVEIVE DPRVVGIWTKNKE\LELSVPKFKIDEFYVDQV
İ					ļ	PPKQVTFAKLNDNIRENFLRDMCKKYGEVEE
						VEILYNPKTKKHLGIAKVVFATVRGAKDAVO
ł	ŀ				ļ	HLHSTSVMGNIHVELDTKGETRMRFYELLV
	1					TGRYTPOTLPVGELDAVSPIVNETLQLSDALK
	ŀ					RLKDGGLSAGCGSGSSSVTPNSGGTPFSODTA
						YSSCRLDTPNSYG/QGTPLTPRLGTPFSQDSSY
	Ì	Ì			Í	SSROPTPSYLFSODPAVTFKARRHESKFTDAY
						NRRHEHHYVHNSPAVTAVAGATAAFRGSSD
						LPFGTVGGTGGSSGPPFKAQPQDSATFAHTPP
						PAQATPAPGFR
1148	2498	A	9302	1026	6	IASIQNADTMPGVGLLVSHFSTLVSRQRCPNY
						ADPQNLTDVSIFLLLEVSGDPELQPVLAGLFL
<u> </u>])				SMCLVTVLGNLLIILAISPDSHLHTPMYFFFSN
				,		LSLPDV\GFTSTTVPK\MIVDI\QSRSRVISYAG
						CLTQKSLFAIFGGTEE\NMLLSVMAYDRFVAI
						CHPLYHSAIMNPCFCAFLVLLSFFFLSLLDSQL
•					i	HSWIVLQFTIIKNVEISNFVCDPSQLLKFACSD
						SIINSIFIYFHKDPERQLVLAGLFLSMCLVTVL
						GNLIIILDVSPDSHLPTPMYFFLSNLSLPDIGFT STTVPKMIVDIQSHGRVIFYAGCLTQMSLFAIF
,						GGMEERHAPECDGL
1149	2499	A	9303	1	699	MASQEKDIFIGWGTIHLFRKPQRSFFGKLLRE
		''		•	0,,	FRLVAADRSMGRYMLFGVINLICTGFLLMWC
						SSINSIALIVSYTYLTIFOLFSLMTCLISYWVTL
		i				RKPSPVYSFGFERLEVLAVFASTVLAQLGALF
1						ILKESAERFLEOPEIHTGRLLVGTFVALCFNLF
						TMLSIRNKPFAYVSEAASTSWLQEHVADLSR
		_	[SLCGIPGLSSIFLPRMNPFVLIDLAGAFALCIT
		L		·		YMLIEI
1150	2500	A	9308	797	693	DRSTSVTRAGVQWCSLGSLQPRTPGLLRSSCL
						SLP
1151	2501	A	9309	205	406	VAIKELPVLWKWSKPTR\TAKEPPQTQQRAG
		1				SKTAAPPCQWSRMASEGPNIPCPGARHSDKQ
						FLICTI
1152	2502	A	9314	913	504	KPSPLITPPAVVLPPSAVLNLVNTFSSFPQVEV
						QGPLCGPRKGRLAVTIPFFGLS/LPKYMDHRR
						PPPHR\EIFFVFLAETGFHRASQAGPDLPTS/S/I
						PPTSA/FPKCWEYRSEPQCLPGCLSFSGILLDL
						GTNVSLRAA
1153	2503	A	9315	392	1	HPHRPRPGFRSPARSSRPCPVLTSLLPPFPSPSP
						PADDLVKAGRDRKDPQVR/ERRLRPNPGRLG
			.			GPR\PRPARARS/CHQPRLTRVCPRSPPPEARA
						PAPAAPARGRGAPKRNRPRTDTRAPRGSSAR
ــــــــــــــــــــــــــــــــــــــ			L			PGNS

SEQ ID NO: of	SEQ ID NO: of	Met hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	""	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
1	ļ	Ì		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
!	}	Ì	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ĺ				peptide	ļ	/=possible nucleotide deletion, \=possible
				sequence		nucleotide insertion
1154	2504	A	9321	331	433	MPCI/QAQYGTPAPSPGPRDHSASDPLTPEFIK
1166	0505		0224	100	000	PT CORPORATION CONTROL TO COMPANY CONTROL TO
1155	2505	A	9324	180	275	MEEPQSDPSVEPPLSQETFSDLWKLLSENNVL
1156	2506	A	9326	383	619	MISPSRTEGDPLPLPP/EGEGQEVRGFGGGPAK
ł	1		ľ	ļ	Ì	EAAQRHCRASVSILRMRRPGQGSSRPARVPL RGPDSHRLREPPPSPP
1157	2507	A	9327	152	292	YERRGRSQGGGSHPAGAQPGGRAIGAGWOS
1137	2307	^	9321	132	232	KEPLWEGLQRSGSPLPG
1158	2508	A	9328	1	430	QELKQGPNPLAPSPSAPSTSAGLGDCNHRVD
			7020	·	'50	LSKTFSVSSALAMLQERRCLYVVLTDSRCFL
ļ				}		VCMCFLTFIQALMVSGYLSSVITTIERRYSLKS
1				1		SESGLLVSCFDIGNLVVVVFVSYFRGRRRRP/
	İ	ĺ	[1	RVAAVGGLLDLEGGEMI
1159	2509	A	9334	108	383	KGNQVNGNGNQLKRKHESMCPVSLTQNTVR
				ł		LMEAGLPQKQAERADELFEAGLVIYVKLDER
			<u> </u>			VLNAL\YSSVGLQWFKESDLSHLRLLEISFR
1160	2510	Α	9338	2	430	FVGRPRGLSDRLEDLFLAGFRVGERLRTAAM
		1	1	ļ		KRYVRILLLGEGAEHVADPVPGGRGVPRGEA
				İ		DHTDQELREEIHKANVERVVHDVSQEATIEKI
[ĺ	RTKWIPLV/RWGDHA/EGPVGIKSYLPSGRSM
1161	2511	A	9341	1	390	EAELPIMSQLTEIETCVEC NSRVDDFVAPGLSEAGKLLGLEFPERORLAA
1101	2311	^	7341	*	390	AVG/CSPMSGVISMSAPFFLGKIIDAIYTNPTV
ĺ	1		ļ			DYSDNLTRLCLGLSGVFLCGAAANAIRVYLM
	1		1			QTSRQRVVKRLRTSLFSSILGQEVAFSDKAGT
j	j					GELI
1162	2512	A	9343	84	837	QGRFRAFCWQRDFLQPPGMRLSALLALASKV
				١.		TLPPHYRYGMSPPGSVADKRKNPPWIRRRPV
	l					VVEPISDEDWYLFCGDTVEILEGKDAGKQGK
		1	1			VVQVIRQRNWVVVGGLNTHYRYIGKTMDYR
ļ		1	l			GTMIPSEAPLLHRQVKLVDPMDRKPTEIEWR
						FTEAGERVRVSTRSGRIIPKPEFPRADGIVPET
	l	ł	i	ļ.	}	WIDGPKDTSVEDALERTYVPCLKTLQEEVME
	İ	l				AMGIKETR\NTRRSIGIEPGAEQLLPNFCPSLE G
1163	2513	A	9346	967	616	DSLALSPRLECSGAISAHCNLTPPGFTPFSCLS
]	1] ~~.]	LPSSWAYRCASPHPDNFFVFLVESGFHHVGO
	1]	AGLKLLISSDPPTSA/FPKCWDYRRD\SSAPAT
l						FSSYQRNNPDLILNDTIMPNIK
1164	2514	A	9347	3	1099	SSFPTCMRTVFHSNTSVSSLLHRPGHVTPQLTI
1	Ì	۱ .		1		HGGWRHHRDHTAIDEWDFNPSKFLIYTCLLL
	1		1			FSVLLPLRLDGIIQWSYWAVFAPIWLWKLLV
L.						VAGASVGAGVWARNPRYRTEGEACVEFKA
	1	1)		MLIAVGIHLLLLMFEVLVCDRVERGTHFWLL
	1					VFMPLFFVSPVSVAACVWGFRHDRSLELEILC
		1				SVNILQFIFIALKLDRIIHWPWLVVFVPLWILM SFLCLVVLYYIVWSLLFLRSLDVVAEQRRTH
	1	ì		l		VTMAISWITIVVPLLTFEVLLVHRLDGHNTFS
	[[1			YVSIFVPLWLSLLTLMATTFRKGGNHWWF
			1			AIRRDF/CODOLPOPTGKPPPPPLTDHHGEKA
		[1			LPLONKDRGSWPASRGSPRLL
1165	2515	A	9362	547	991	DVSIGPPLLRRPCSGREQTRSLSFPSDPESSFSP
]	1	1			VPEGVRLADGPGHCKGRVEVKHQNQWYTV
		1				CQTGWSLRAAKVVCRQLRCGRAVLT\QKRC
	1		ļ			TKHAYGRKPIWLSQMACSGPEPTLHDCPFRP
		L	L			LGEDTLFHVEYTSVHGRERLSAKD
1166	2516	A	9363	201	387	PPILRWTPPSGKNFFFFFFFESEFY/SSPRVECS
1167	2517	A	9368	707	1007	GAISAHLAHCNLCLPGSSDSPASAFQVAS
110/	2517	<u> </u>	7300	707	1087	AVLTPCLSPCSPSRIPRP\SRPYPGRRSLSHTPP

SEQ ID NO: of No: of	ine, ne, n, nd, le PDLLRPT LPDPPPW T WMLKGS HDPNAID SRVLVGS GELGGRA AALGPSL SGVNPN VMAYWS HGPVGH LLADDDA	
nucleotide sequence USSN 09496 uence UEncure USSN 09496 uence UENcure USSN 09496 uence UENcure USSN 09496 uence UENcure USSN 09496 uence UENcure USSN 09496 uence UENcure USSN 09496 uence UENcure UENcu	ne, n, n, nle PDLLRPT LPDPPPW T WMLKGS HDPNAID SRVLVGS GELGGRA AALGPSL SGVNPN VMAYWS HGPVGH LLADDDA	
cotide sequence when the sequence of peptide sequence when the seq	ne, n, n, nle PDLLRPT LPDPPPW T WMLKGS HDPNAID SRVLVGS GELGGRA AALGPSL SGVNPN VMAYWS HGPVGH LLADDDA	
sequence uence	n, on, on, on, on, on, on, on, on, on, o	
uence 914 ng to first amino acid residue of peptide sequence of peptide sequence sequence	n, on, on, on, on, on, on, on, on, on, o	
amino acid residue of peptide sequence T=Threonine, V=Valine, W=Tryptopha sequence	IN, IN letion, \text{\text{\$\text{\$-}\sigma}} possible nucleotide deletion, \text{\text{\$\text{\$-}\sigma}} possible nucleotide deletion, \text{\text{\$\text{\$-}\sigma}} possible nucleotide deletion, \text{\$\text{\$\text{\$-}\sigma}} possible nucleotide deletion, \text{\$\text{\$\text{\$-}\sigma}} possible nucleotide deletion, \text{\$\text{\$\text{\$\text{\$-}\sigma}} possible nucleotide insertion properties of the nucleotide insertion properties of the nucleotide insertion properties of the nucleotide deletion, \text{\$\text{\$\text{\$\text{\$-}\sigma}} possible nucleotide deletion, \text{\$\text{\$\text{\$\text{\$-}\sigma}} possible nucleotide insertion properties of the nucle	IN, IN IN IN IN IN IN IN IN IN IN IN IN IN
Peptide sequence	PDLLRPT LPDPPPW T WMLKGS HDPNAID SRVLVGS GELGGRA AALGPSL SGVNPN VMAYWS HGPVGH LLADDDA	
PRPLILYAPAPRPAGTAFIPHSHPPP ATPA/TPCPS1PPPPRPHPT/OPSTAI PLPFPPPSS/RPPPPDCSTSYSPTPPP ATPA/TPCPS1PPPPRPHPT/OPSTAI PLPFPPPSS/RPPPPDCSTSYSPTPPP TPPFPPPSS/RPPRPDCSTSYSPTPPP MMLSEETSAVRPQKQTRFNGAKLV PITVTSAVIIVLMLIMM/FSPWLAT LTARLLPPSAAHWFGTDEVGRDLF; QQSILAGLVVATTGMIGSPLECLF DAIFMRVMDIMRS/IPSLVLTMEKT, FNAMQASSEH DAIFMRVMDIMRS/IPSLVLTMEKT, FNAMQASSEH LVFVLLFIFAKRQIMRFAMKSLRGP NAPKDLKEEIDILLSRVHNIK YEPH NSARRMEAMASGSNWLSGVNVVL LVFVLLFIFAKRQIMRFAMKSLRGP NAPKDLKEEIDILLSRVHNIK YEPH ILLLTICAAGIGGTFQFGYNLSIINAF TNETWQARTGEPLPDHLVLLMWSL GGLFGALLAGPLAITLGRKKSLLW AAILFGFSRKAGSFEMIMLGRLASW SMNIQFMLPGGESAPKELRGAVAM LGIVMGQVVGLSTTAATGLRGLAC ERAACQGCRARRPWELFQHRALRR VLGSAMELCGNDSVYA YASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTTPSFFALNQFTLQKKKK WGGTTPSFFALNQFTLQKKKK TPDKRKGLAY/IQQTDDSLIHFCWK EDDLIIFPDDCEFKRPQCPNGRVY SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETKK SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETKK CONNGRED CO	LPDPPPW T WMLKGS HDPNAID SRVLVGS GELGGRA AALGPSL SGVNPN VMAYWS HGPVGH LLADDDA	
ATPA/TPCPSLPPPPRPLHPTQPSTAI PLPFPPPSSKPPRPDCSTSYSPTFPPP 1168 2518 A 9375 511 15 MMLSEETSAVRPQKQTRFNGAKLV PITVTSAVIIVLMLIMM/iFSPWLAT LTARLLPPSAAHWFGTDEVGRDLF-QQSILAGLVVVATTGMIGSPLECLF DAIFMRVMDIMRS/IPSLVLTMEKT.FNAMQASSEH GNGRVAPRDPGAVASAEPGLTTHD NSARRMEAMASGSNWLSGVNVVLLVFVLLFIFAKRQIMRFAMKSLRGP NAPKDLKEEIDILLSRVHNIK YEPH NSARRMEAMASGSNWLSGVNVVLLVFVLLFIFAKRQIMRFAMKSLRGP NAPKDLKEEIDILLSRVHNIK YEPH ILLLTICAAGIGGTFQFGYNLSIINAF TNETWQARTGEPLPDHLVLLMWSI GGLFGALLAGPLAITLGRKKSLLVVAAAIL.FGFSKKAGSFEMIMGRLASW SMNIQPMLPGGESAPKELRGAVAM LGIVMGQVVGLSTTAATGLRGILAG ERAACQGCRARRPWELFQHRALRR VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNOFTLQKKKK LFPSLLPCSWSTSNKYLAEFRAGKM TPDKRKGLAJVQQTDDSLIHFCWK EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETK	LPDPPPW T WMLKGS HDPNAID SRVLVGS GELGGRA AALGPSL SGVNPN VMAYWS HGPVGH LLADDDA	
PLPFPPPSS/RPPRDCSTSYSPTFPPF	TWMLKGS HDPNAID SRVLVGS GELGGRA AALGPSL SGVNPN VMAYWS HGPVGH LLADDDA	
1168	WMLKGS HDPNAID SRVLVGS GELGGRA AALGPSL SGVNPN VMAYWS HGPVGH LLADDDA	
PITVTSAVIIVLMLLMMIFSPWLAT LTARLLPPSAAHWFGTDEVGRDLF: QQSLAGLVVVATTGMIGSPLECLF DAIFMRYMDIMRS/IPSLVLTMEKT. FNAMQASSEH 1169 2519 A 9377 42 410 GNGRVAPRDPGAVASAEPGLTTHD NSARRMEAMASGSNWLSGVNVVL LVFVLLFIFAKRQIMRFAMKSLRGP NAPKDLKEEIDILLSRVHNIKYEPH 1170 2520 A 9378 302 1303 GVSGFSASVLRQRRMEDELEPSLRF ILLLTICAAGIGGTFQFGYNLSIINAF TNETWQARTGEPLPDHLVLLMWSI GGLFGALLAGPLAITLGRKKSLLIVI AAILFGFSRKAGSFEMIMLGRLASW SMNIQPMLPGGESAPKELRGAVAN LGIVMGQVVGLSTTAATGLRGLAC ERAACQGCRARRPWELFQHRALRR VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNKYLAEFRAGKM TPDKRKGLAY/IQQTDDSLIHFCWK EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK	HDPNAID SRVLVGS GELGGRA AALGPSL SGVNPN VMAYWS HGPVGH LLADDDA	
LTARLLPPSAAHWFGTDEVGRDLF: QQSILAGLVVVATTGMIGSPLECLF DAIFMRVMDIMRS/IPSLVLTMEKT. FNAMQASSEH 1169 2519 A 9377 42 410 GNGRVAPRDPGAVASAEPGLTTHD NSARRMEAMASGSNWLSGVNVVL LVFVLLFIFAKRQIMRFAMKSLRGP NAPKDLKEEIDILLSRVHNIKYEPH 1170 2520 A 9378 302 1303 GVSGFSASVLRQRMEDELEPSLRF ILLLTICAAGIGGTFQFGYNLSIINAF TNETWQARTGEPLPDHLVLLMWSI GGLFGALLAGPLAITLGRKKSLLWI AAILFGFSRKAGSFEMIMLGRLASW SMNIQPMLPGGESAPKELRGAVAN LGIVMGQVVGLSTTAATGLRGLAG ERAACQGCRARRPWELFQHRALRR VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLIPCSWSTSNKYLAEFRAGKM TPDKRKGLAY/IQQTDDSLIHFCWK EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK	SRVLVGS GELGGRA AALGPSL SGVNPN VMAYWS HGPVGH LLADDDA	
QQSILAGLVVVATTGMIGSPLECLF DAIFMRVMDIMRS/IPSLVLTMEKTA FNAMQASSEH 1169 2519 A 9377 42 410 GNGRVAPRDPGAVASAEPGLTTHD NSARMEAMASGSNWLSGVNVVD LVFVLLFIFAKRQIMRFAMKSLRGP NAPKDLKEEIDILLSRVHNIK YEPPH NAPKDLKEEIDILLSRVHNIK YEPPH ILLLTICAAGIGGTFQFGYNLSIINAF TNETWQARTGEPLPDHLVLLMWSI GGLFGALLAGPLAITLGRKKSLLVVI AAILFGFSRKAGSFEMIMLGRLASW SMNIQPMLPGGESAPKELRGAVAN LGIVMGQVVGLSTTAATGLRGLAC ERAACQGCRARRPWELFQHRALRR VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNK YLAEFRAGKM TPDKRKGLAY/IQQTDDSLIHFCWK EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK	GELGGRA AALGPSL SGVNPN VMAYWS HGPVGH LLADDDA	
DAIFMRVMDIMRS/IPSLVLTMEKTA FNAMQASSEH 1169 2519 A 9377 42 410 GNGRVAPRDPGAVASAEPGLTTHD NSARRMEAMASGSNWLSGVNVVL LVFVLLFIFAKRQIMRFAMKSLRGP NAPKDLKEEIDILLSRVHNIK YEPYH NAPKDLKEEIDILLSRVHNIK YEPYH ILLLTICAAGIGGTFQFGYNLSIINAF TNETWQARTGEPLPDHLVLLMWSI GGLFGALLAGPLAITLGRKKSLLVVI AAILFGFSRKAGSFEMIMLGRLASW SMNIQPMLPGGESAPKELRGAVAN LGIVMGQVVGLSTTAATGLRGLAC ERAACQGCRARRPWELFQHRALRR VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNK YLAEFRAGKM TPDKRKGLAY/IQQTDDSLIHFCWK EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK	SGVNPN VMAYWS HGPVGH LLADDDA	
FNAMQASSEH	SGVNPN VMAYWS HGPVGH LLADDDA	
1170 2521 A 9381 2 412 RGPASAQEDERATTAPLERVRARG 1170 2520 A 9384 20 355 GWNGRSTEASPAAEAPHYPHKETIK	VMAYWS HGPVGH LLADDDA	
NSARRMEAMASGSNWLSGVNVVL LVFVLLFIFAKRQIMRFAMKSLRGP NAPKDLKEEIDILLSRVHNIKYEPH 1170 2520 A 9378 302 1303 GVSGFSASVLRQRRMEDELEPSLRF ILLLTICAAGIGGTFQFGYNLSIINAF TNETWQARTGEPLPDHLVLLMWSI GGLFGALLAGPLAITLGRKKSLL\VI AAILFGFSRKAGSFEMIMLGRLASW SMNIQPMLPGGESAPKELRGAVAM LGIVMGQVVGLSTTAATGLRGLAC ERAACQGCRARRPWELFQHRALRR VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNKYLAEFRAGKM TPDKRKGLAYJQQTDDSLIHFCWK EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK	VMAYWS HGPVGH LLADDDA	
1170 2520 A 9378 302 1303 GVSGFSASVLRQRRMEDELEPSLRE ILLLTICAAGIGGTFQFGYNLSIINAF TNETWQARTGEPLPDHLVLLMWSJ GGLFGALLAGPLAITLGRKKSLL\VI AAILFGFSRKAGSFEMIMLGRLASW SMNIQPMLPGGESAPKELRGAVAM LGIVMGQVVGLSTTAATGLRGLAC ERAACQGCRARRPWELFQHRALRE VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNKYLAEFRAGKM TPDKRKGLAYJQQTDDSLIHFCWK EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK	HGPVGH LLADDDA	
1170 2520 A 9378 302 1303 GVSGFSASVLRQRRMEDELEPSLRF ILLLTICAAGIGGTFQFGYNLSIINAF TNETWQARTGEPLPDHLVLLMWSI GGLFGALLAGPLAITLGRKKSLL\NY AAILFGFSRKAGSFEMIMLGRLASW SMNIQPMLPGGESAPKELRGAVAN LGIVMGQVVGLSTTAATGLRGL\AC ERAACQGCRARRPWELFQHRALRR VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNKYLAEFRAGKM TPDKRKGLAYJQQTDDSLIHFCWK EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK	LLADDDA	
1170 2520 A 9378 302 1303 GVSGFSASVLRQRRMEDELEPSLRF II.LLTICAAGIGGTFQFGYNLSIINAF TNETWQARTGEPLPDHLVLLMWSI GGLFGALLAGPLAITLGRKKSLL\VI AAILFGFSRKAGSFEMIMLGRLASW SMNIQPMLPGGESAPKELRGAVAN LGIVMGQVVGLSTTAATGLRGL\AC ERAACQGCRARRPWELFQHRALRR VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNKYLAEFRAGKM TPDKRKGLAYJQQTDDSLIHFCWKI EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK	RTOIOGR	
ILLLTICAAGIGGTFQFGYNLSIINAF TNETWQARTGEPLPDHLVLLMWSI GGLFGALLAGPLAITLGRKSLLIVN AAILFGFSRKAGSFEMIMLGRLASW SMNIQPMLPGGESAPKELRGAVAN LGIVMGQVVGLSTTAATGLRGLAQ ERAACQGCRARPWELFQHRALRR VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNKYLAEFRAGKM TPDKRKGLAY/IQQTDDSLIHFCWK EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK	ADOLO	
TNETWQARTGEPLPDHLVLLMWSI GGLFGALLAGPLAITLGRKKSLL\VI AAILFGFSRKAGSFEMIMLGRLASW SMNIQPMLPGGESAPKELRGAVAN LGIVMGQVVGLSTTAATGLRGLAC ERAACQGCRARRPWELFQHRALRR VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNK YLAEFRAGKM TPDKRKGLAY/IQQTDDSLIHFCWKI EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK		
GGLFGALLAGPLAITLGRKKSLL\VI AAILFGFSRKAGSFEMIMLGRLASW SMNIDP\MLPGGESAPKELRGAVAM LGIVMGQVVGLSTTAATGLRGLAC ERAACQGCRARRPWELFQHRALRR VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNK YLAEFRAGKM TPDKRKGLAY/IQQTDDSLIHFCWK EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK	IVSLVPI	
AAILFGFSRKAGSFEMIMLGRLASW SMNIQPMLPGGESAPKELRGAVAM LGIVMGQVVGLSTTAATGLRGLAC ERAACQGCRARRPWELFQHRALRR VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNKYLAEFRAGKM TPDKRKGLAYAQQTDDSLIHFCWK EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK		
SMNIQP\MLPGGESAPKELRGAVAM LGIVMGQVVGLSTTAATGLRGLAC ERAACQGCRARRPWELFQHRALRR VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNKYLAEFRAGKM TPDKRKGLAYAQQTDDSLIHFCWK EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK		
ERAACQGCRARRPWELFQHRALRR VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNKYLAEFRAGKM TPDKRKGLAYNQQTDDSLIHFCWKI EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK		
VLGSAMELCGNDSVYAYASSVFRK KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNKYLAEFRAGKM TPDKRKGLAY/IQQTDDSLIHFCWKI EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK	ELEELEE	
KIQYAIIGTGSCELLTAVVSVSLEGA WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNK.YLAEFRAGKM TPDKRKGLAY/IQQTDDSLIHFCWK/ EDDLLIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK	QVTSLV	
WGGTPRSFALNQFTLQKKKK 1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNK YLAEFRAGKM TPDKRKGLAY/IQQTDDSLIHFCWKI EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETVK	AGVPEA	
1171 2521 A 9381 2 412 RGPASAQEDERARTAPLERVRARGI LFPSLLPCSWSTSNK YLAEFRAGKM TPDKRKGLAY/IQQTDDSLIHFCWKI EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETIK	LPPPAL	
LFPSLLPCSWSTSNKYLAEFRAGKM TPDKRKGLAY/IQQTDDSLIHFCWKI EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETVK		
TPDKRKGLAY/JQQTDDSLIHFCWKJ EDDLIIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETVK		
EDDLIFPDDCEFKRLPQCPNGRVYV SKRLFFWMQEP 1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETVK		
SKRLFFWMQEP		
1172 2522 A 9384 20 355 GWNGRSTEASPAAEAPHVPHKETNK	LKIKAG	
on to the last the la	AAMGTO	
CTHGGKVRPDPHDMLTTVVHKIKL		
LQLCAIMISDYLKSSIYTVEKRLGLF		
ASFNEVGNTALIVLESY	100	
1173 2523 A 9393 430 87 LCQCIVPGQQKETFSLNPSSATVRFY	'L*LSLO	
QRKEDQ+III.*YHLNKDCLHIFMSAI		
KIFVLFDFNIMFETPFYII*FIFLFSQN		
IRPPISFSKINNGP		
1174 2524 A 9397 77 374 ERLEIGRLGGERGSGPASCLRVIDVS		
RLVKLALLQLLRAFYGIKVKGVRVI		
ESSSTLIRVS*FGVPCNALAHFGVTH	F*YILDF	
LGML LGML LGML 1175 2525 A 9399 66 397 HESSRADRDKMDTRGSTYTDADRY	W100==	
1 Indicate Distriction 1		
AKMNKWSKGKVRDKLNNLVLFDT.		
CKEVPNYKLITLAVVSERLKIPGSLA	KAALHE	
1176 2526 A 9408 2 299 LDLTHVLSLSISLTVTILGTTFGMVIP	I Dian	
GERGYAQNGDF*DAQLDDYSFSCYS GAPNSLTRAYDDP*VKISGLECQKV		
KCLNL		
1177 2527 A 9416 2 402 CNFLRSSRIRVHSTPAASTMPPKVDF	W - L +	
YLRCTGGEVRATSALAPKIGPLGLSS		
FV*ATGDWNVLIISVILTIRILLSHIFV	NEIKVV	
DHLIAFWDLQSLIFLHVIFSLFITLLLI	NEIKVV SIKVGVD	
1178 2528 A 9419 142 426 TPLFDLWPRVVLSWLETVLTSLRTR	NEIKVV SIKVGVD VPPFFCF	
ACRIMPTTVDDVLEHGGEVHFLOKO	NEIKVV SIKVGVD VPPFFCF FCFFSIF	
ALI*DTFAPIYVGIVFLGFTPDHRCRS	NEIKVV SIKVGVD VPPFFCF CFFSIF RAASGPP	
1179 2529 A 9420 1450 1655 LSSAGTKMNLN*KNYWPGASAHAC	NEIKVV SIKVGVD VPPFFCF FCFFSIF RAASGPP OMLYLL	
GQSRCITRSGDRDHPG*HGETPSVLK	NEIKVV SIKVGVD VPPFFCF FCFFSIF RAASGPP OMLYLL PGVAEL	
WWRAP .	NEIKVV SIKVGVD VPPFFCF FCFFSIF RAASGPP MLYLL PGVAEL NPSTLG	

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	l	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline,
seq-	uence	1	09/496	correspondi	to last amino acid residue	O=Glutamine, R=Arginine, S=Serine,
uence			914	ng to first amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
	ł			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
			ļ	peptide	sequence	/=possible nucleotide deletion, \=possible
ì	l	i	l	sequence	1	nucleotide insertion
1180	2530	A	9422	176	375	HRPQTTRPDWKPRT*PQGK*GRLSSEISPASPP
1100	2330	^	7422	170	373	SRFSRSTKPVPPKADPPARQKLTGVLHAPLLK
l		İ	1			L
1181	2531	A	9436	2	274	PIAASLRMYNLOPYTEENLICTAFATMVETVP
						IARTILDRLTGIPHGYCFVE*ADWATADKCVH
{	ĺ	1		ſ	1	IYNGKPLPGATPLLSLQLHQLAHLGS
1182	2532	A	9442	3	240	VDKCSSKSIVLSEYCPHCMCSLSTDPKPFGQL
					1	SMILK*MGAGDEKISAMGKARVDHRELYLGL
		ĺ	l			LYPTEDYKLTFRARH
1183	2533	Α	9444	384	3	LKDFQPWALHDWPLFCCCTFLLFLVLECFTR
						KGCSGWAPWLSLQCQHFGRPRWADHLRSGV
	i	1	l			RDQPGQYSKTTFLPKIQKLAGHSGAHL*S*LL
	•	1				ERMRWKNRLNPGGRSCSEPRWHHCTPGWAT
	ļ.,,,,	<u> </u>	<u> </u>		<u> </u>	ERG
1184	2534	A	9462	391	655	LSGFKSLMPKIPLQYIYVRVRTTWSFCLPLDG
	ľ	ĺ	1	1	ľ	RKLMLS*YSK*LT*KYNILPEYSRMTLPPGMV
		<u> </u>			ļ <u>.</u>	IHTCNPSTLGGRAGWIV*AQEFET
1185	2535	A	9467	215	566	RCPMWQGQASRMDPAKAKDREASTCCSLA
		1		ì	ļ	WWWGWECWVRALKLSSGPAGPLACWVAK KKSLSLSGPVYPSEKGAGLYVF*DRVSLCHPG
		1				1
1106	0526	 	0460	275	460	WSAVVQFWLTAASNSCFSLLSSWDYRCA HIPOLHTKTHYVPTRMVNKI*QIDNSKPWQR
1186	2536	Α	9468	275	452	GG*TGILTHCW*ESKLVQPLWKIVWHYQ
1187	2537	A	9469	388	3	EVAPGPSQILPRRVTDGGDRPQFSLPGPRLPQ
1107	2557	Α	9409	300	3	SSRGAEPCLSNCIHSPAPRKQRMGDSDQ*STP
		ł	İ			NPASPHPEAPOEPWDSASGSVGSFSLGRGAK
1	1			'		ASS*VPGKGRGPRQGSELLAETILELFLALAN
1	1	1	1)		S
1188	2538	A	9471	124	397	TMDKKNRHGNSLDMASEIHMTGPMCLIENTT
****						GRLMANPEALKILSAITQPMVEEAIAGLYRAC
1	ŀ	1	1			*FYLTNNLAGMKKGLCLGSTEQAHTIGI
1189	2539	A	9480	584	769	GHVQSQHFGRPRRADHLRSGDRDHPG*HDET
	l	1		1		PSLLKIQKISWAWWRAPVVPATWEAEAEEW
1	ì			<u> </u>		R
1190	2540	A	9483	463	86	VTVGLTLLLRGAPRFTAG*PPSGGGPPLAPLL
1		1				PRQHCTLQTHRHLHPEAPVKV*KT*RLFPGLR
	1	1				GASSCRRRCNPVLAARKAGSPRSHSTRENC
			1			RRSRCPDTAHRRRRRGRRRNPSCVRSPRWR
1191	2541	A	9489	1	411	LADALCLSAAATGAVRPGARAQPSTRRLSP
1	1	l ·	1	1		SVRVCCRAAAASNLLYSSCLQRHSERASEEG
						ERGSLSAKCCSLVLRGGCSSSNSHSFRRIT*EI
1			1	1		MAAFVLLSYEQRPLKRPRLGPPDVYPPDPKQ KEEELTAVNVK
1102	25/2	 	0407	290	161	VSFLSMSSGHCIRSTRGSKMVSWSVIAKIQEI*
1192	2542	A	9497	389	101	CEEDERKMAREFLAEFMSTYVMMNIHMIVE
1	1					KDTYSDHEEINTS
1193	2543	A	9509	186	1	IAKSQ*KRWQRSGAMETLKHGWWECKLVQF
1133	2343	^	לטכל	100	1 *	FGKTFVNVN*S*TYVYPCDKIILLLGLYPTEM
1194	2544	HA-	9512	58	433	PLQRSKCLTLRCLRAKPWAWSQSPRACSSAL
1124	2344	^	7312	1 30	"55	LKSSRSRASSLNVQCILQSNPQGHQRI*KQKA
	1	1		1	1	SSKGQQFRR*KEHPFMLKTLNKLRIEGT*LKI
1	}	1	1		1	RRAIYDNPTANIIVEGQKLEAFPLRTGTRQ
1195	2545	A	9515	595	1223	GHGAPSFQTQVPRTP*ASWPVVPAASESAPAP
1,1,3	1	1	12,5	1		AGGGASLPVAAGSCAAAPHTEPGAPQHLLDC
	1	1		1	1	PCPLCLARPPRRPLPDTCYGPGSGRSASLAEPP
		1	1			LPRCSCAPLRSASAPOVS*CV*AVNLLPHNL*
1		1	1	1	1	PLHLLLHD*EKAWGFLFSSASHCFQGQICLLP
1		1	1			APGSGPCGATARPSRGGRAGGSRARRPIPPGP
		1	1		1	GTRRTPSGCQNPAASGG
						

SEO ID	SEQ ID	Mct	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	1	110u	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
	peptide					
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
Ì	· '		Ì	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
ļ		i		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
i	,			peptide		/=possible nucleotide delction, \=possible
1)	l	}	sequence	ļ	nucleotide insertion
1196	2546	Α	9518	229	468	RSPTATPAPHAMGPGAPFARGGRPLPLLGAM
	1					AERVAPGWDLHTPYLPRTNSRRTPHL**EPHA
1	Į į	l	ł		1	GYIGALFPMSGGWPGGQ
1197	2547	A	9521	289	448	IAWLSGLFFPSNQANLCFLCYKLTADSRYRG
''''	2347	^)))]]	267	1770	HAMRHLTGNTSMAIRFL*ADSRFOVORARYE
1						
1100	0540		0504			APNWKYKYGY*IPVDMLC
1198	2548	Α	9524	204	1	KNKKTTKCLSIVTLNISGPNQ*NKRHRVAEWI
	i	ŀ				VKQEPNICHL*ETHFPFRETYRLKEREQKKRK
<u></u>	L					SSYS
1199	2549	Λ	9546	1785	1943	GGRFKESKLTNAGWQRNSFFIGPPKSIPWAA
1		1	İ		1	V*QRGDGKNPGVTHLNRPVGTX
1200	2550	A	9548	186	1	VNAEKEF*KIQHYFMTKSQNKLHIEHTYLKPI
1	1	1			1	KAIYDKWTSDIMLNLOKL*AFFLRVIVROI
1201	2551	A	9549	591	2	SSVVEFPRGPRSSLPPLDSTFPCGSSPNWTGGC
1.201		^	''''	".	~	GSCPSGE*LVSPGSEQRKKYSNSNVIMHETSO
	1		l			YHVOHLATFIMDKSEAITSVDDAIRKLVOLSS
			l	ł	Į.	
1	1		1	1		KEKIWTQEMLLQVNDQSLRLLDIESQEELEDF
1	!					PLPTVQRSQTVLNQLRYPSVLLLVCQDSEQSK
1						PDVHFFHCDEVEAELVHEYMESALTDCRLGK
[Ĺ		L	l	<u> </u>	AMRP
1202	2552	A	9552	428	1	KYGNEGHWSRQCPNPGKPIRPCPLCRGPHWK
						LDCERPPQGPLPSLPELAKTSYSDLTGLATED
]	1				1	*WGPGMDAPATTIASSKTRVTLMVAGRPVFF
Į.	ł				1	LI*YRATYSALPNFSGPTQSSQVSVVGIDGQV
l				1		SKPRATPPLFCSLHTF
1203	2553	A	9568	517.	738	RRKFERKOKO*RYREGKOYRORDKMKEWG
1203	2333	**	7500	J.,	1 .30	EKEKRREKGEREERKMRHRERKGESGORD
					1	TMENWRVERLTEKER
1204	2554	A	9573	83	416	
1204	2334	A	93/3	83	415	EDKRLRLVDGDSRCAGRV*IYHDGFWGTICD
ł	1 .			1	ł	DGWDLSDAHVVCQKLGCGVAFNATVSAHFG
						EGSGPIWLDDLNCTGTESHLWQCPSRGWGQ
						HDCRHKEDAGVICSEFTALR
1205	2555	. A	9577	64	424	ARGSCPTRPRTANGRMGETKDAPQMLVTFK
ļ.						DVAVTFFREEWRQLVLVHRTLYR*GMLETC
i	[*	GLLDTLRHNVPQPDVVHLLYHGTQLLIVKRE
						VSHSPCAGDMRELFTREATLTPHPYNNGA
1206	2556	A	9584	38	476	TLGAVLFSEVSKESSTSHSGGQLGRQNRHPKL
]						SNFITPSSPRLKP*TASSQRNLGQILNMFLTAV
1						NPQPLSTPSWQIETKYSTKVLTGNWMEERRK
ł						GLPYKHLITHHQEPPHRYLISTYDDHYNRHG
}						YNPGLPPLRTWNGQKLLWL
1207	2557	A	0506		412	
120/	ادىء	A	9586	2	412	LRSSPAALLRALCITTVTGTALALRSRVATTN
ł			}			PDGCRNVLRPKYYRLCDKAESWGIALETVPT
]	1	•				GVAVTSWAIMLTVLTLVCKGQDYNRRQKLP
1			1			THILCLL*EKGIFGLTFAFIIGLDGSTGPTRFFL
	L		L		l	FGILFSICFS
1208	2558	Α	9597	122	3	IKNYWPGMVAHACNPSPLGGRGRWIA*AQK
1	1	1				FADAWADAW
1209	2559	A	9611	148	558	KSLRNVWDLLNNTWKADRFFCHSSRTSTIRK
1		"				GDPGPTFSKMSIWTSGRTSSSYRHDEKRNIYQ
						RIRDHDLLDKRKTVTALKAGEDRAILLGLAM
İ		}]			MVCSIMM*FLLGITLLRSYMOSVWTRESQCT
l		1				
1010	05/0		0610	204		LLNASITETFNC
1210	2560	A	9618	384	2	SLHDMLMLAEQQQKQKWAVNTQNTAWSNA
l]			DSKFGQRILEKMEWSKGRGLGVQEQGGPDDI
]]	Ì	J			KVQVKNNDLGLQATINNEANWIAHQDDFNW
1	1	l	1	I	l	LLAELNTCQRQETADS***WSPKNSHVGKDS
1						28,22,1104,1421,22
						GELSAK
1211	2561	A	9620	316	610	

CEO ID	SEQ ID	Met	Lego	Dandinkad	1 Dec State of and	14
SEQ ID NO: of	NO: of	hod	SEQ ID NO:	Predicted beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	1100	in in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
		Į		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
	ł	ł		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide		/=possible nucleotide deletion, \=possible
	1	ļ		sequence	1	nucleotide insertion
		 		 		LGRAWWLTPVIPTLWEAKAGGSPE*D*AGRG
		İ				GSRL*SQHFGRPRRVDHLRSAVQDQPGQHGE
	İ	ļ		}	1	TPSLLKIQKIN*VWGRRL*SSYSEAEAGESL
1212	2562	A	9623	297	344	QFPVDGDYQKIEKITQLFQAQNLSLCLAMTR
		İ				TREL*KGGGKGRHE*AVVPFLKKGGYGVKAP
		ł			1	AILNTSNCT*CF*ETKMLSDDPKACVFEVSSA
	l					DL*NTSFGVIR
1213	2563	A	9624	2	356	AELSLASTACGRNTSGDSLPDYDRAPISSPLA
						TSGTILSAISCLWDLPTPVLRVGLSCQPSMSSQ
		1				IPRMYSTDVEAAVNSLEDLYLQAYYAYLCVG
			L			LYFHRDDMALEGVSRFL*ELAE
1214	2564	A	9634	776	912	SLSRWVRAKL*VPYNQENCLNPRGGGCSEPR
			<u> </u>			SHYCTPAWATEKDS
1215	2565	Α	9636	220	426	KPGNFAVSSEY*DITSGQLKTAVRG*IEMTST
						EENFGEKLHDIGFGNGFLDKT*KAQATKAKI
	224	<u> </u>	 			DK
1216	2566	A	9637	391	76	CFLEDGCTQAS*AEEAAVSPSMAEEEQGSTSC
						RERRSIRFKMKNHSPDDTIKENVTISNIRTRKI
			j			NHLPETERNLLEHGLMYIRLNAAFCSLVAHS
1217	2567	<u> </u>	9655	2000	0420	LFGFILKAT
1217	2307	A	9000	2008	2432	LHCKMGALETQTHPCSQNMLRSLQKCCCKV
						EEHHLQPVQVLQTLLHSATAGTGCRRPARPP
						PAPPTPTPWRSRQSGKQSERAS*LKGRGRYGL
	ļ	j		•		GALGGRGGRALGGSRWPPPLPGETLFSGCKH RRRRGSDAAPGEEAGT
1218	2568	Ā	9658	3	405	HASARALLSPNLSPNNKMAISGGPVLGFFIIA
	2500	^	7030	J ,	403	VLMSAQEPWAIKEEHVIIQAEFYLNPDOSGEF
	1	}		·		MLDFEGEDTFHGDMAKKETVWRLE*LARLD
				•		NFEAQRALANIAADQAALEIMDMGSDYTLIP
		ŀ				NVPPKVTVL
1219	2569	A	9662	3	284	PDWTEKRKMQDTGSILPLHWFGFGYAALVA
-						YGGIIGYVKAGSVPSLAAGLLFGSLSGLGAYQ
	ļ		ļ			LSQDPRNVWVFLATSGTLAGIMGMRFYHSG
						KL
1220	2570	Α	9669	200	699	LLLTGYIQTLQNQQLSGNQQEMQAVDNLTSA
		ļ				PGNTSLCTRDYKITQVLFPLLYTVLFFVGLITN
			:			GLAMRIFFQIRSKSNFIIFLKNTVISDLLMILTF
			1		i	PFKILSDAKLGTGPLRTFVCQVTSVIFYFTMYI
			}			SISFLGLITIDRYQKTTRPFKTSNPKNLLGAKIL
						<u>K</u>
1221	2571	A	9676	164	562	KERDSSTFSAAMTTMQGMEQAMPGAGPGVP
	[QLGNMAVIHSHLWKGLQEKFLKGEPKVLGV
			4			VQILTALMSLSMGITMMCMASNTYGSNPISV
	ĺ		1			YIGYTIWGSVMFIISGSLSIAAGIRTTKGLVRG
1000	0500		0.000			SLGMNITSS
1222	2572	A	9688	43	412	VAKMVKCCSAIGCASRCLPNSKLKGLTFHVF
				i	i i	PTDENIKRKWVLAMKRLDVNAAGIWEPKKG
						DVLCSRHFKKTDFDRSAPNIKLKPGVIPSIFDS
1002	2572		000	200		PYHLQGKREKLHCRKNFTLKTVPATNYNH
1223	2573	A	9696	308	564	RTSMGILYSEPICQAAYQNDFGQVWRWVKE
		İ) i		ļ	DSSYANVQDGFNGDTPLICACRRGHVRIVSFL
1224	2574	- <u>-</u>	0700		(22	LKKECLCQPQKPERENLLALCCE
1224	ا 14 دک	A	9700	3	632	DAWASGGELGSLFDHHVQRAVCDTRAKYRE
]		 	GRRPRAVKVYTINLESQYLLIQGVPAVGVMK
				ſ		ELVERFALYGAIEQYNALDEYPAEDFTEVYLI
			ļ	l		KFMNLQSARTAKRKMDEQSFFGGLLHVCYA
				1		PEFETVEETRKKLQMRKAYVVKTTENKDHY
				!	' l	VTKKKLVTEHKDTEDFRQDFHSEMSGFCKA
	لـــــــــــــــــــــــــــــــــــــ					ALNTSAGNSNPYLPYSCELPLCYFSSK

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide]	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		Į	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
				amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
				peptide	•	/=possible nucleotide deletion, \=possible
1				sequence	İ	nucleotide insertion
1225	2575	Α	9710	1	163	RSGCVLRMTEWETGAPAVAETPDIKLFGKWS
1			ļ			TDDVHINDISLQDYIAGVRLILL
1226	2576	Α	9713	82	492	QGLPSFLPAFGPSGSWLGPAPTLGSSCNTVDT
.		1				ICHGYSEIRPLFYLSFCDLLLGLCWLTETLLYG
1		1				ASVANKDIICYNLQAVGQIFYISSFLYTVNYI
i		l				WYLYTELRMKHTQSGQSTSPLVIDYTCRVCQ
1 1		{	i			MAFVFSSLI
1227	2577	A	9720	3	416	GKWKRTQVPLLGEECADMDLARKEFLRGNG
		1			`**	LAAGKMNISIDI.DTNYAFI.VLNVGRVTLGEN
!		ł	1			NRKKMKDCQLRKQQNENVSRAVCALLNSGG
1 1	,	l				GVIKAEVENKGYSYKKDGIGLDLENSFSNML
1		İ	ĺ			PFVPNFLDFMQNGNYF
1228	2578	A	9723	278	411	EASSSNTVASNVADKTDPHSMNSRVFIGNLN
		, · ·				TLVLQKSDVEAVF
1229	2579	A	9725	121	902	LFAMSGFENLNTDFYQTSYSIDDQSQQSYDY
		 ^^	7,23		-02	GGSGGPYSKQYAGYDYSQQGRFVPPDMMQP
						QQPYTGQIYQPTQAYTPASPQPFYGNNFEDEP
					:	PLLEELGINFDHIWOKTLTVLHPLKVADGSIM
i i				•		NETDLAGPMVFCLAFGATLLLAGKIOFGYVY
1 1						GISAIGCLGMFCLLNLMSMTGVSFGCVASVL
			}			GYCLLPMILLSSFAVIFSLQGMVGIILTAGIIG
						WCSFSASKIFISALAMEGQQLLVAYPCALLYG
1 1					'	VFALISVF
1230	2580	Α	9739	11 .	247	TFVLNMNTPKEEFQDWPIVRIAAHLPDLIVYG
1.250	2500	11	7,37	• •	271	HFSPERPFMDYFDGVLMFVDISGKCKRDVCL
						MWMSNRLAWEFTCRA
1231	2581	A	9744	37	1100	TPLFDFWPGFVLSWLQPLSASLRARRAASGPP
'23'	2361	A.	7/7	37	1100	ACRIMPTTVDDVLEHGGEFHFFQKQMFFLLA
				•		LLSATFAPIYVGIVFLGFTPDHRCRSPGVAELS
				_		LRCGWSPAEELNYTVPGPGPAGEASPROCRR
1 1						YEVDWNQSTFDCVDPLASLDTNRSRLPLGPC
						RDGWVYETPGSSIVTEFNLVCANSWMLDLFO
						SSVNVGFFIGSMSIGYIADRFGRKLCLLTTVLI
1 1						NAAAGVLMAISPTYTWMLIFRLIQGLVSKAG
1 1			•			WLIGYILITEFVGRRYRRTVGIFYOVAYTVGL
						LVLAGVAYALPHWRWLQFTVALPNFFFLLY
1 1						YWCIPESPRWLISONKNAEAMRIIKHIAKKNG
					i	KSLPASL
1232	2582	A	9753	164	517	PGPGMQGPPPITPTSWSLPPWRAYVAAAVLC
		·- I	7.00		· · · ·	YINLLNYMNWFIIAGVLLDIQEVFQISDNHAG
]						LLQTVFVSCLLLSAPVFGYLGDRHSRKATMS
<u> </u>			.			FGILLWSGAGLSSSFISPRYSWLF
1233	2583	A	9757	25	419	LPAPWTERVRKSEGLVGTCLGDPMASPRTVT
		· ·	7,57	~	-	IVALSVALGLFFVFMGTIKLTPRLSKDAYSEM
				İ		KRAYKSYVRALPLLKKMGINSILLRKSIGALE
]						VACGIVMTLVPGRPKDVANFFLLLLVLAVLF
					ļ	FHOLV
1234	2584	Α	9765	71	456	RLELDWGFSLHFLPVAYLCPLSSGFEMNVQP
"""	2704	•		′¹	420	
]						CSRCGYGVYPAEKISCIDQIWHKACFHCEVC KMMLSVNNFVSHOKKPYCHAHNPKNNTFTS
	[l	[
						VYHTPLNLNVRTFPEAISGIHDQEDGEQCKSV
1235	2585	A	9767		550	FHWD
1233	دەرى	^	3101	52	559	IRSGAMSVDKAELCGSLLTWLQTFHVPSPCA
1		1				SPQDLSSGLAVAYVLNQIDPSWFNEAWLQGI
						SEDPGPNWKLKVTSGLLIRGQTGEEMTRDGP
	ł		ļ		ŀ	ARHMSWVMGRKRDRCLVINHLFIHSSMEYSP
			j			CARPGHSARNNTDKNLPHTAILLVTSNTYTTI
1226	2592	~ —	0770	262		KINFQAGRSGSCL
1236	2586	<u>A</u>	9770	352	608	FRGEALTVRFLTKRFIGEYASNFESIYKKHLC

SEO ID	SEO ID	Met	SEO	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide	"""	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	O=Glutamine, R=Arginine, S=Serine.
"""	ľ	l	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
		l		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
1	1	1		peptide	Sequence	/=possible nucleotide deletion, \=possible
}	ł	ł	ł	sequence		nucleotide insertion
	+	 	 	sequence	<u> </u>	LERKQLNLEIYDPCSQTQKAKFSLTSELHWA
]					DGFVIVYDISDRSSFAFAKALI
1237	2587	A	9793	266	515	NILAIIYFPFPRLFLLRDSQSNPKAFALTLCHH
1231	2307	Α	5173	200] 313	QKIKNFQILPVSIDALTPPLVVCFLVSFLTHFS
		l				RYKPTRPVCITQFQGCS
1238	2588	A	9802	537	967	
1238	2388	Ι ^	9802	337	90/	ELGAGRSDREAMEAAVKEEISVEDEAVDKNI
	Ì	{				FRDCNKIAFYRRQKQWLSKKSTYRALLDSVT
		ŀ	ł			TDEDSTRFQIINEASKVPLLAEIYGIEGNIFRLK
			ŀ			INEETPLKPRFEVPDVLTSKPSTVRLISCSGDT
						GSLILADGKGDLKC
1239	2589	Α	9805	105	540	VPGDPAMVRAGAVGAHLPASGLDIFGDLKK
		ļ	1			MNKRQLYYQVLNFAMIVSSALMIWKGLIVLT
	1	1	1		1	GSESPIVVVLSGSMEPAFHRGDLLFLTNFRED
	1	1	ļ)	PIRAGEIVVFKVEGRDIPIVHRVIKVHEKDNG
]	· '	DIKFLTKGDNNEGDDRGSYK
1240	2590	A	9819	3	305	TDGRDPLPCAARRRGGGGECCGAGWVAEWS
		}		1		PQPLDPAMLLWMQGFVLEAVACQDNDDYLR
	1			[YGILFEDLDCNGDGVVDIIELQEGLRNWSSAF
	1	j		j		DPNSEEHG
1241	2591	A	9834	841	1209	SPARGKSNRTDVMITAPKNKKMTENLAAPEA
	1					LDSSTHSSSTATQSRAKMNTPAPTPSTVPAIPR
	į	l				GGSGGPPPCAPHDRVSSVLQCDTQAMDHKTE
	İ	i			İ	SSHSVVEFLFKRTKTPSPFHPAVRENRN
1242	2592	A	9843	3 .	589	TISCGPATEPPASLLSSASSDDFCKEKTEDRYS
		l	1 20.5] ` `	307	LGSSLDSGMRTPLCRICFQGPEQGELLSPCRC
		İ				DGSVKCTHQPCLIKWISERGCWSCELCYYKY
				,		HVIAISTKNPLQWQAISLTVIEKVQVAAAILGS
	1	ĺ				LFLIASISWLIWSTFSPSARWQRQDLLFQICYG
	1			•		MYGFMDVMIVAVDSEDMVQAAKEVGKRWS
	1					DIPP
1243	2593	A	9846	198	411	WRISHHAGKMPVMKGLLAPONTFLDTIATRF
	1 2373	^	7040	176	711	DGTHSNFILANAQVAKGFPIVYCSDGFCELAG
						FARTEVMO
1244	2594	A	9848	116	650	PICGFLYLCSAMASESSPLLAYRLLGEEGVAL
1277	2354	^	7040	110	630	
	i .		1			PANGAGGPGGASARKLSTFLGVVVPTVLSMF
						SIVVFLRIGFVVGHAGLLQALAMLLVAYFILA
	1					LTVLSVCAIATNGAVQGGGAYCILQHRWTG
	i					VWPVLPAREVMISRTLGPEVGGSIGLMFYLA
1245	2505	<u> </u>	0040		1600	NVCGCAVSLLGLVESVLDVFGA
1245	2595	Α	9849	573	1620	KSKCRFPEGLSEGFGPMRKEALSSGSVQEAE
		l				AMLDEPQEQAEGSLTVYVISEHSSLLPQDMM
	1	1	} ;			SYIGPKRTAVVRGIMHREAFNIIGRRIVQVAQ
			[_ [AMSLTEDVLAAALADHLPEDKWSAEKRRPL
						KSSLGYEITFSLLNPDPKSHDVYWDIEGAVRR
	1			į		YVQPFLNALGAAGNFSVDSQILYYAMLGVNP
						RFDSASSSYYLDMHSLPHVINPVESRLGSSAA
	}	1				SLYPVLNFLLYVPELAHSPLYIQDKDGAPVAT
						NAFHSPRWGGIMVYNVDSKTYNASVLPVRV
					'	EVDMVRVMEVFLAQLRLLFGIAQPQLPPKCL
]			LSGPTSEGLMTWELDRLLWARSVENLATATT
						TLTSLA
1246	2596	A	9850	114	464	PPQLGAQRVREPRHPDVRAPLRVTSPGLRSRS
						ARSLGRRPRIAMVTVGNYCEAEGPVGPAWM
				ļ		QDGLSPCFFFTLVPSTRMALGTLALVLALPCK
]			RRERPAGADSLSWGAGPRISSYV
1247	1-2	A	9851	2	327	FVRNKKMTRSCSAVGCSTRDTVLSRERGLSF
	2597					- 12111411141114001 TOUING I VISITATION
	2597	^	7031			
	2597	Α	7031	1		HQFPTDTIQRSKWIRAVNRVDPRSKKIWIPGP
•	2597	Λ	7031			

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion
1248	2598	A	9853	58	444	RVDDFVYSKGKDAGGADVSLACRRQSIPEE FRGITVVELIKKEGSTLGLTISGGTDKDGKPR VSNLRPGGLAARSDLLNIGDYIRSVNGIHLTR LRHDEIITLLKNVGERVVLEVEYELPPPGGCP WT
1249	2599	A	9856	2	1265	LPPPRPSRHRRGRAGTRASAAAAAGPTVSAV RAPVRGQDSGAGTPQGRLAGRGAHLSRVGA SGSGVAAGPAARIIAPRRRCADAGEAVGASC GRCAVALLSGVCTLVSTHVCVGSGCPGAAGT PMGAGDAGASAESAVTTAPQEPPARPLQAGS GAGPAPGRAMRSTTILALLALVLLYLVSGAL VFRALEQPHEQQAQRELGEVREKFLRAHPCV SDQELGLLIKEVADALGGGADPETNSTSNSSH SAWDLGSAFFFSGTIITTIGGGGDWHVGGGK ELPHGGRCRETEGSQVAPRLPASPLCPGYGN VALRTDAGRLFCIFYALVGIPLFGILLAGVGD RLGSSLRHGIGHIEAIFLKWHVPELVRVLSA MLFLLIGCLLFVLTPTFVFCYMEDWSKLEAIY FVIVTLTTVGFGDYVA
1250	2600	Α	9873	2	652	FVVPSPCGGIPGRAPNGASRPTMGNSASRNDF EWVYTDQPHTQRRKEILAKYPAIKALMRPDP RLKWAVLVLVLVQMLACWLVRGLAWRWLL FWAYAFGGCVNHSLTLAHDISHNAAFGTGR AARNRWLAVFANLPEGVPYAASFKKYHVDH HRYLGGDGLDVDVPTRLEGWFFCTPARKLL WLVLQPFFYSLRPLCVHPKAVTRMEVLNTLV OLA
1251	2601	A	9875	150	1209	PVIMPLHFSPGDIVRPSCCVSSSPKLRRNAHSR LESYRPDTDLSREDTGCNLQHISDRENIDDLN MEFNPSDHPRASTIFLSKSQTDVREKRKSLFIN HHPPGQIARKYSSCSTIFLDDSTVSQPNLKYTI KCVALAIYYHIKNRDPDGRMLLDIFDENLHPL SKSEVPPDYDKHNPEQKQIYRFVRTLFSAAQL TAECAIVTLVYLERLLTYAEIDCPANWKRIV LGAILLASKVWDDQAVWNVDYCQILKDITVE DMNELERQFLELLQFNINVPSSVYAKYYFDL RSLAEANNLSFPLEPLSRERAHKLEAISRLCED KYKDLRRSARKRSASADNLTLPRWSPAIIS
1252	2602	A	9879	6	376	KRPDSRPPAQYRAGPTRPRTRGCELLYWKAT KAVGIKMGSLSTANVEFCLDVFKELNSNNIG DNIFFSSLSLLYALSMVLLGARGETEEQLEKV WNSSEVCSEPRSLSCSRSGSAKLILSLYQ
1253	2603	A	9880	180	388	KEQAELLYGLYCQCDLTLSSHPSSVPAMSSC NFTHATFVLIGIPGLEKAHFWVGFPLLSMYVA AMFGNC
1254	2604	A	9881	19	494	VISFQIITDTIMDSSTAHSPVFLVFPPEITASEYE STELSATIFSTQSPLQKLFARKMKILGTIQILF GIMTFSFGVIFLFTLLKPYPRFPFIFLSGYPFWG SVLFINSGAFLIAVKRKTTETLIILSRIMNFLSA LGAIAGIILLTFEFHPRSKLHL
1255	2605	A	9896	72	386	RPGREQRDCFQAPPLGLGGRQTDMMHHPLT GATCVGLPNVGMCPQLSGALTFMYLQQGNQ EATVAPDTMAQPYASAQFAPPQNGIPGEYTA PHPHPAPEYTGOTT
1256	2606	A	9902	95	399	SGGPAGLLHRPVLPKMGLSGLLPILVPFILLG DIQEPGHAEGILGKPCPKIKVECEVEEIDQCTK PRDCPENMKCCPFSRGKKCLDFRKVSLTLYH KEELE
1257	2607	A	9905	374	459	EHLKSTPNRLGVVAHTCNPSTLGGRGGW

CEO ID	SEQ ID	Met	CEO	Dundintal	Dendists J J	I Amino ocid money (A - Al-el- C C
SEQ ID NO: of	NO: of	hod	SEQ ID NO:	Predicted beginning	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
mucl-	peptide	"tou	in NO:	nucleotide	nucleotide location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	1	1	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
	1		1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
1	1	}	1	residue of	sequence.	Y=Tyrosine, X=Unknown, *=Stop codon.
-	İ	1		peptide	- Coquesto.	/=possible nucleotide deletion, \=possible
i	1	1	1	sequence	}	nucleotide insertion
1258	2608	A	9911	364	1974	AGPGVPAVGGRWASGPGLGGRTLCSGPPDH
1	1					QRRGPSCGASGDPQCVGSPHPQRARPLLARP
j	į	ļ		j	J	GARLLPGHLPSPRPPRLPTGQPPAAAFRGPVR
1	l			1		PQGGGHIHPLPTPGGRPCFAVSEGSGSALLLS
				1	Į.	YLGECGSSSYVTGAACISPVLRCREWFEAGLP
	1	1	1	•		WPYERGFLLHQKIALSRYATALEDTVDTSRL
1	1	1				FRSRSLREFEEALFCHTKSFPISWDAYWDRND
					[PLRDVDEAAVPVLCICSADDPVCGPPDITTLTT
1	}		1			ELFHSNPYFFLLLSRHGGHCGFLRQEPLPAWS
			i			HEVILESFRALTEFFRTEERIKGLSRHRASFLG
		j	j]]	GRRRGGALQRREVSSSSNLEEIFNWKRSYTRL
			1			MAAAAGAAAAPGSREPQDRPECGAGHPGPR
		1				YYRHPERWLLRPEAFLGPLRTRAPSAEDSQR
1					1	ERPAARSGPEMRVRYPVVAAVLAPYLALSQD
1	1			[ì	PMVKSSASGQGASGSYNHVREEMLIKAGGA
1			1			MSRRVVRQSKFRHVFGQAAKADQAYEDIRV
		i	1	i	İ	SKVTWDSSFCAVNPKFLAIIVEAGGGGAFIVL
1259	2609	A	9919	693	935	PLAK
1239	2009	A	1 9919	093	933	GCFKFIGESTCCWIFPSSVTTQCVVAKAPRAA
					· ·	TLSKAERLRSQPGPEQGGSSYRPRTPTAAAIL
1260	2610	A	9921	455	1082	PPRPGRSHRKRKLVSTK
.200	2010	Α .	7721	433	1002	QRSCLCSAIEKDGGDVKALYRRSQALEKLGR LDQAVLDLQRCVSLEPKNKVFOEALRNIGGO
1	1	1		İ		IQEKVRYMSSTDAKVEQMFQILLDPEEKGTE
		1				KKQKASQNLVVLAREDAGAEKIFRSNGVOLL
1					i '	QRLLDMGETDLMLAALRTLVGICSEHQSRTV
1		ŀ	}	1 .		ATLSILGTRRVVSILGVESQAVSLAACHLLQV
1						MFDALKEGVKKGFRGKEGAIIV
1261	2611	A	9928	1	438	GFRGAEAPGAAQAPKKKKPRPTEGGPGAGSG
	1	1				RGKDPYRGPTLLHQPKPPKDEFLSSLESYEIAF
1	ļ	l	1			PTRVDHNGALLAFSPPPPQRQRRGTGATAES
1				1		RLFYKEASPSTHFLLNLTRSSRLLAGHVSVEY
L	<u> </u>		1)	1	WTREGLAWQRADRPHCLYA
1262	2612	A	9931	168	435	AAEMGRAGAAAVIPGLALLWAVGLGGPPPA
ł	}	ł	ł			PPRLPFCLQELQGRHALHTFSLERTCSYQDFL
			1			WADEGRLLHVGAQDLATWHTLSPLGLW
1263	2613	A	9938	247	488	RMSATSVDQRPKGQGNKVSVQNGSIHQKDG
]	1]	1			CNDDDFEPYLRSPDNQSNSYPPMSDPYMPGY
1021		!	<u> </u>	L.,		YAPSIGFPYSLGEAAWSQL
1264	2614	A	9941	61	277	ESIGLTALGPRRRPWEHRWSDPITLKMKGWG
1	ļ					WLALLLGALLGTAWARRSQDLHCGACKAVR
1965	2615	ļ	0055	<u> </u>		RRVRQFNIYDY
1265	2615	A.	9956	2	522	FVASEVSKMPVPASWPHPPGPFLLLTLLLGLT
1	J]	}			EVAGEEELQMIQPEKLLLVTVGKTATLHCTV
	[.				ļ .	TSLLPVGPVLWFRGVGPGRELIYNQKEGHFP
]	}	J]	ļ		RVTTVSDLTKRNNMDFSIRISSITPADVGTYY
1].			į.	ļ [CVKFRKGSPDHVEFKSGAGTELSVRGEYSVG
1266	2616	A	10002	242	207	FLSQVWWWLSSHPFMN
1200	2010	^	10002	243	387	PKNNACHLLFTAVCQPRCKHGECIGPNKCKC
1267	2617	<u> </u>	10004	36	707	HPGYAGKTCNQGRKTV
120/	201/	^	10004	30	707	LPAPASTWSVARETMASSSVPPATVSAATAG
1	i					PGPGFGFASKTKKKHFVQQKVKVFRAADPLV
						GVFLWGVAHSINELSQVPPPVMLLPDDFKAS
]	1	[SKIKVNNHLFHRENLPSHFKFKEYCPQVFRNL
	ŀ	1				RDRFGIDDQDYLVSLTRNPPSESEGSDGRFLIS YDRTLVIKEVSSEDIADMHSNLSNYHQVRPLS
1		1	1	ľ		SPILSLSSLLTYSSAIVSNRCQLGRKLIGRENP
1268	2618	A	10005	2	209 ·	GEGYELFVPSNGVPAVCHMVGRRPHRAVLSP
				-	207	SQDELEHSLGESAAQGAAGVVLWVSWENTR
		Ь	لـــــــــــــــــــــــــــــــــــــ			-455FFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF

SEQ ID	SEQ ID	Met	SEQ	Predicted	D121	14
NO: of	NO: of	hod	ID NO:	beginning	Predicted end nucleotide	Amino acid sequence (A=Alanine C=Cysteine,
nucl-	peptide	1100	in NO.	nucleotide	location	D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	1	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	1	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	20.700		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
l ucinco		1	714	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
1		ŀ		residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ł				peptide	l codamen	/=possible nucleotide deletion, \=possible
}		ļ		sequence		nucleotide insertion
 						TKVSLGLA
1269	2619	A	10010	245	688	FGMLKNKGHSSKKDNLAVNAVALODHILHD
]]	1	"""	LQLRNLSVADHSKTQVQKKENKSLKRDTKAI
		l	l			IDTGLKKTTQCPKLEDSEKEYVLDPKPPPLTL
İ		1				AQKLGLIGPPPPPLSSDEWEKVKQRSLLQGDS
}	1	ļ			ļ	VQPCPICKEEFELRPQVFSIRG
1270	2620	A	10011	2	588	RVDDFVRPLPPGLMSRSRASIHRGSIPAMSYA
	i	ļ	1	ł		PFRDVRGPSTHRTQYVHSPYDRPGWNPRFCII
				İ		SGNQLLMLDEDEIHPLLIRDRRSESSRNKLLR
			1			RTVSVPVEGRPHGEHEYHLGRSRRKSVPGGK
		l		1		QYSMEGAPAAPFRPSQGFLSRRLKSSIKRTKS
		•		İ		QPKLDRTSSFRQILPRFRSADHDRYRGWSMW
		1		j		DEIDV
1271	2621	A	10013	209	363	LPAPPNLSPRLSFGFQFPGGNDNYLTITGPSHP
				ļ		FLSGAEVSQSCRRRGGRA
1272	2622	A	10014	7	388	SAVTISWKWRSVMGIQTSPALLASLGAGLVT
	}	}	1			LLGLAVGSYLVRRSRRPQVTLLDPNEKDLLR
		ļ .	j			LIDKTLSARSPCKHIYLSTRIDGSLSIRPYTPVT
						SDEDQGYVDIDIKVYLKGVHPTFPEGGKMSH
1273	2623	Α	10016	1	1339	MAARTLGRGVGRLLGSLRGLSGQPARPPCGV
						SAPRRAASGPSGSAPAVAAAAAQPGSYPALS
		ł				AQAAREPAAFWGPLARDTLVWDTPYHTVW
						DCDFSTGKIGWFLGGQLNVSVNCLDQHVRKS
						PESVALIWERDEPGTEVRITYRELLETTCRLA
				,		NTLKRHGVHRGDRVAIYMPVSPLAVAAMLA
						CARIGAVHTVIFAGFSAESLAGRINDAKCKVV
						ITFNQGLRGGRVVELKKIVDEAVKHCPTVQH
						VLVAHRTDNKVHMGDLDVPLEQEMAKEDP
				,		VCAPESMGSEDMLFMLYTSGSTGMPKGIVHT
						QAGYLLYAALTHKLVFDHQPGDIFGCVADIG
						WITGHSYVVYGPLCNGATSVLFESTPVYPNA
						GRYWETVERLKINQFYGAPTAVRLLLKYGD
					[AWVKKYDRSSLRTLGSVGEPINCEAWEWLH
	0.00					RVVGDSRCTLVDTWWQT
1274	2624	Α	10017	1	3750	FRPQGTPRSPASHVLTMSAPDEGRRDPPKPKG
1				. 	ì	KTLGSFFGSLPGFSSARNLVANAHSSARARPA
1					ļ	ADPTGAPAAEAAQPQAQVAAHPEQTAPWTE
						KELQPSEKMVSGAKDLVCSKMSRAKDAVSS
- 1			' l		İ	GVASVVDVAKGVVQGGLDTTRSALTGTKEV
						VSSGVTGAMDMAKGAVQGGLDTSKAVLTG
			-			TKDTVSTGLTGAVNVAKGTVQAGVDTTKTV
İ	ľ	{	1		Į.	LTGTKDTVTTGVMGAVNLAKGTVQTGVETS
]				KAVLTGTKDAVSTGLTGAVNVARGSIQTGV
						DTSKTVLTGTKDTVCSGVTGAMNVAKGTIQT
				i	j	GVDTSKTVLTGTKDTVCSGVTGAMNVAKGT
1				ļ	j	IQTGVDTSKTVLTGTKDTVCSGVTGAMNVA
1		[ĺ	. [KGTIQTGVDTTKTVLTGTKNTVCSGVTGAVN
i			1			LAKEAIQGGLDTTKSMVMGTKDTMSTGLTG
ļ						AANVAKGAMQTGLNTTQNIATGTKDTVCSG
1		1	1	i		VTGAMNLARGTIQTGVDTTKIVLTGTKDTVC
1		1	1	ļ	İ	SGVTGAANVAKGAVQGGLDTTKSVLTGTKD
į				.		AVSTGLTGAVNVAKGTVQTGVDTTKTVLTG
ł	- 1	1	1	ł		TKDTVCSGVTSAVNVAKGAVQGGLDTTKSV
						VIGTKDTMSTGLTGAANVAKGAVQTGVDTA
	ļ		1		ļ	MAN TOTAL TOTAL PROPERTY AND AND AND AND AND AND AND AND AND AND
f						KTVLTGTKDTVTTGLVGAVNVAKGTVQTGM
						DTTKTVLTGTKDTIYSGVTSAVNVAKGAVQT
						DTTKTVLTGTKDTIYSGVTSAVNVAKGAVQT GLKTTQNIATGTKNTFGSGVTSAVNVAKGAA
						DTTKTVLTGTKDTIYSGVTSAVNVAKGAVQT

SEQ ID NO: of nucl- cotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion VAKGAIQGGLDTTKSVLTGTKDAVSTGLTGA VKLAKGTVQTGMDTTKTVLTGTKDAVCSGV
						TGAANVAKGAVQMGVDTAKTVLTGTKDTV CSGVTGAANVAKGAVQTGLKTTQNIATGTK NTLGSGVTGAAKVAKGAVQGGLDTTKSVLT GTKDAVSTGLTGAVNLAKGTVQTGVDTSKT VLTGTKDTVCSGVTGAVNVAKGTVQTGVDT AKTVLSGAKDAVTTGVTGAVNVAKGTVQTG VDASKAVLMGTKDTVFSGVTGAMSMAKGA VQGGLDTTKTVLTGTKDAVSAGLMGSGNVA TGATHTGLSTFQNWLPSTPATSWGGLTSSRT TDNGGEQTALSPQEAPFSGISTPPDVLSVGPEP AWEAAATTKGLATDVATFTQGAAPGREDTG LLATTHGPEEAPRLAMLQNELEGLGDIFHPM NAEEQAQLAASQPGPKVLSAEQGSYFVRLGD LGPSFRQRAFEHAVSHLQHGQFQARDTLAQL QDCFRL
1275	2625	A	10025	124	415	TILARKKEKTCPCKKEIGRNSRSGMYSRKAM YKRKYSAANTKVEKKKKEKVLAPVTKPVGG DKNGGTRVVKLPTMPRYYPTEDVPRKLLSHG KKPFS
1276	2626	Ā	10030	3	507	GGSLRFSPPRVPSCSRVFCPVPPGGCGLPSPMS ASRPQSPTTPWCLPRRYMKHKRDDGPEKQED EAVDVTPVMTCVFVVMCCSMLVLLYYFYDL LVYVVIGIFCLASATGLYSCLAPCVRRLPFGK CRIPNNSLPYFHKRPQARMLLLALFCVAVSV VWGVFRNEDQ
1277	2627	Α	10035	51	869	YSRFTVPLPATMASSEVARHLLFQSHMATKT TCMSSQGSDDEQIKRENIRSLTMSGHVGFESL PDQLVNRSIQQGFCFNILCVGETGIGKSTLIDT LFNTNFEDYESSHFCPNVKLKAQTYELQESN VQLKLTIVNTVGFGDQINKEERQLGRSQSTEN PQKYRSEQHPVEPKKCTSFWKGALGKWAGIE SSGQSAQQPYLPINSPPHRLADVADVHLFSSV LSGAFGCYHLDVTVNEFKKQQNRDEQEGYS KGDQEQGSWKHGADPLRGGEM
1278	2628	A	10036	3	457	RAFDVRRKKSLRPCCPRDFHAGCLTVSGPST VMGAVGESLSVQCRYEEKYKTFNKYWCRQP CLPIWHEMVETGGSEGVVRSDQVIITDHPGDL TFTVTLENLTADDAGKYRCGIATILQEDGLSG FLPDPFFQVQVLVSSASSTENSVKTP
1279	2629	A	10039	214	435	NDSLVPMSSWRSCARAPSSESAWRRSAATRR SRKCLRTKRKRWSSGKGTQMQSTLSETPRRA QMPCMWWYPFWG
1280	2630	A	10043	2	344	RATWHNAGKEREAVQLMAGAEKRVKASHS FLRGLFGGNTRIEEACEMYTRAANMFKMAK NWSAAGNAFCQAAKLHMQLQSKHDSATSFV DAGNAYKKADPQGKTARHVACYLCV
1281	2631	A	10080	620	818	VIYKLDSSLFSYFIYFFIFETESHFLPLMKWTG PIMAHCSLKILASRNSADSAFLSAGDTSLSHST
1282	2632	A	10084	3	1640	SASIIIRGDKRASGEVGIAPSSRHILIGEPSAKY NGTAIISLVRGPGILGEVTVFWRIFPPSVGEFA ETSGKLTMRDEQSAVIVVIQALNDDIPEEKSF YEFQLTAVSEGGVLSESSSTANITVVASDSPY GRFAFSHEQLRVSEAQRVNITIIRSSGDFGHVR LWYKTMSGTAEAGLDFVPAAGELLFEAGEM RKSLHVEILDDDYPEGPEEFSLTITKVELQGR GYDFTIQENGLQIDQPPEIGNISIVRIIIMKNDN AEGIIEFDPKYTAFEVEEDVGLIMIPVVRLHGT

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutaminc, R=Argininc, S=Serinc, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion YGYVTADFISQSSSASPGGVDYILHGSTVTFQ HGQNLSFINISIIDDNESEFEEPIEILLTGATGG AVLGRHLVSRIIIAKSDSPFGVIRFLNQSKISIA NPNSTMILSLVLERTGGLLGEIQVNWETVGPN SQEALLPQNRDIADPVSGLFYFGEGEGGGVRTILLTTYPHEEIEVEETFIIKLHLVKGEAKLDSRAK DVTLTIQEFGDPNGVVQFAPETLSKKTYSEPL ALEGPLLITFFVRRVKGTFGEIM
1283	2633	A	10088	316	516	MGSKTLPAPVPIHPSLQLTNYSFLQAVNGLPT VPSDHLPNLYGFSALHAVHLHQWTLGYPAM
1284	2634	A	10091	2	569	HLXRS FVSPSRAMASALIYVSKFKSFVILVVTPLLLP LVILMPAKFVRCAYVIILMAIYWCTEVIPLAV TSLMPVLLFPLFQILDSRQVCVQYMKDTNML FLGGLIVAVAVERWNLHKRIALRTLLWVGA KPARLMLGFMGVTALLSMWISNTATTAMMV PIVEAILQQMEATSAATEAGLELVDKGKAKE LP
1285	2635	A	10092	290	728	KQSTRPDVMTLYPLHWQEEMSGESVVSSAVP AAATRTTSFKGTSPSSKYVKLNVGGALYYTT MQTLTKQDTMLKAMFSGRMEVLTDSEGWIL IDRCGKHFGTILNYLRDGAVPLPESRREIEELL AEAKYYLVQGLVEECQAALQV
1286	2636	Α	10100	1	574	RPRGRGAWAGPGGDYSGVRRQQRRRTRISGS QRGSDAAGTMGCCTGRCSLICLCALQLVSAL ERQIFDFLGFQWAPILGNFLHIIVVILGLFGTIQ YRPRYIMVYTVWTALWVTWNVFIICFYLEVG GLSKDTDLMTFNISVHRSWWREHGPGCVRR VLPPSAHGMMDDYTYVSVTGCIVDFQYLEVI HSA
1287	2637	Α	10103	252	376	RSRMGDKPIWEQIGSSFIQHYYQLFDNDRTQL GAIYVSFQL
1288	2638	A	10107	1	478	MEEDESRGKTEESGEDRGDGPPDRDPTLSPS AFILRAIQQAVGSSLQGDLPNDKDGSRCHGL RWRRCRSPRSEPRSQESGGTDTATVLDMATD SFLAGLVSVLDPPDTWVPSRLDLRPGESEDM LELVAEVRIGDRDPIPLPVPSLLPRLRAWRTG KT
1289	2639	A	10113	237	438	LLSRMPSTNRAGSLKDPEIAELFFKEDPEKLFT DLREIGHGSFGAAYFARDVRTNEVVAIKKMS YSG
1290	2640	Α	10114	367	856	RGAKAKSAVLPPGPPCSSILILSPPAPLTPRSPG TEATRPTAMSKSLKKKSHWTSKVHESVIGRN PEGQLGFELKGGAENGQFPYLGEVKPGKVAY ESGSKLVSEELLLEVNETPVAGLTIRDVLAVI KHCKDPLRLKCVKQGESSGLLSVLPGGGTAR GAGQ
1291	2641	A	10116	128	591	RTIRETERRSALSCSVLKSEPLPGLQPQASQQR RRRLPGRRQVQVQEGGGSGLRAWVLAMASV LGSGRGSGGLSSQLKCKSKRRRRRRSKRKDK VSILSTFLAPFKHLSPGITNTEDDDTLSTSSAE VKENRNVGNLAARPPPSGDRARGGATR
1292	2642	A	10121	1	749	QRRRFRAGLWGGHGLTDGLRRNGGCGCSAR VPRVGERLRGHRCPDPLCLLLDMLFLSFHAG SWESWCCCCLIPADRPWDRGQHWQLEMADT RSVHETRFEAAVKVIQSLPKNGSFQPTNEMM LKFYSFYKQATEGPCKLSRPGFWDPIGRYKW DAWSSLGDMTKEEAMIAYVEEMKKIIETMP MTEKVEELLRVIGPFYEIVEDKKSGRSSDITSD

SEO II	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of		hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-	1	USSN	location	corresponding	l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence		ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
1	1	ļ		amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan.
		Ī	1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
ì	i	}	1	peptide	}	/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
			i	1.	 	LGNVLTSTPNAKTVNGKAESSDSGAESEEEE
1		1	j		1	AC
1293	2643	A	10124	2	989	PLMSLVRVVEFVAASSAQKTPSRLENYYMVC
1	1	-	*****	1	1 303	KADEKFNQLVHFLRNHKQEKHLVFFRYSSGL
	İ		ŀ			CGRGIRDSARMCSTCACVEYYGKALEVLVK
1	Ī	1		•	1	GVKIMCIHGKMKYKRNKIFMEFRKLQSGILV
ı	1	l		1		CTDVMARGIDIPEVNWVLQYDPPSNASAFVH
	ı	ļ				RCGRTARIGHGGSALVFLLPMEESYINFLAIN
	İ	İ			1	QKCPLQEMKPQRNTADLLPKLKSMALADRA
		ļ]	J	VFEKGMKAFVSYVQAYAKHECNLIFRLKDL
	j	ļ				DFASLARGFALLRMPKMPELRGKQFPDFVPV
1	į	}				DVNTDTIPFKDKIREKQRQKLLEQQRREKTEN
i		i	·		İ	EGRRKFIKNKAWSKOKAKKK
1294	2644	A	10129	91	1042	VTMYKDCIESTGDYFLLCDAEGPWGIILESLA
]		}				ILGIVVTILLLLAFLFLMRKIQDCSQWNVLPTQ
1		l				LLFLLSVLGLFGLAFAFIIELNQQTAPVRYFLF
1						GVLFALCFSCLLAHASNLVKLVRGCVSFSWT
1		l		1		TILCIAIGCSLLQIIIATEYVTLIMTRGMMFVN
1	1	'				MTPCQLNVDFVVLLVYVLFLMALTFFVSKAT
1						FCGPCENWKQHGRLIFITVLFSIIIWVVWISML
1						LRGNPQFQRQPQWDDPVVCIALVTNAWVFL
	1	Ì	l			LLYIVPELCILYRSCRQECPLQGNACPVTAYQ
İ	1		٠			HSFQVENQELSRDKWKVLLNSDFLSHSGA
1295	2645	Α	10133	376	518	RPRVVTHNSQWCFLPQDHPGWLPGQSGAPG
1				,		GRGAPRQEGPGSSWRQV
1296	2646	A	10135	3	551	EWSLDPFMGIMSGQVGDLSPSQEKSLAQFRE
1				·		NIQDVLSALPNPDDYFLLRWLQARSFDLQKS
						EDMLRKHMEFRKQQDLANILAWQPPEVVRL
ł				•		YNANGICGHDGEGSPVWYHIVGSQDPKGLLL
1	į į					SASKQELLRDSFRSCELLLRECELQSQKLGKR
						VEKIIAIFGLEGLGLRDLWKPGIELLQE
1297	2647	A	10138	48	407	MVSSCCGSVCSDQGCGQDLCQETCCRPSCCE
						TTCCRTTCCRPSCCVSSCCRPQCCQSVCCQPT
						CSRPSCCQTTCCRTTCYRPSCCVSSCCRPQCC
L	l. i	- 1	1			QPVCCQPTCCRPSCCETTCCHPXCC
1298	2648	A	10156	94	453	GGNRKSAEMFSQVPRTPASGCYYLNSMTPEG
	1 1		[Ī		QEMYLRFDQTTRRSPYRMSRILARHQLVTKI
			l			QQEIEAKEACDWLRAAGFPQYAQLYEDSQFP
L	<u> </u>			į		INIVAVKNDHDFLEKDLGEPLCRRLNT
1299	2649	A	10161	1	393	PRFSELVDGRGRVSARFGGSPSKAATVRSQPT
			J			ASAQLENMEEAPKRVSLALOLPEHGSKDIGN
	1 1		- 1			VPGNCSENPCQNGGTCVPGADAHSCDCGPGF
	1 1	. 1	- 1	!		KGRRCELACIKVSRPCTRLFSETKAFPVWEGG
	1			ļ		VCHHV
1300	2650	A	10162	98	391	AKIASLERIMPANYTCTRPDGDNTDFRYFIYA
		ļ		l		VTYTGILGPGLIGNILALWVFYGYMKETKRA
	1 1	- 1	i			VIFMINLAIADLLQVLSLPI.RIFYYLKHDWPF
<u></u>	1 1	ł]		1	VPV
1301	2651	A	10165	1	7545	PGIRVGITSQTGLSSNLQENCSKLAFISSHGTE
				- }		KQLQCMPMEGRGRASSSISDLQGKGFEKGTG
	[l	1	ļ	İ	EKHVPGVGSARHSPQASAGGSPWQRGKAQT
		,	Į	I		RWLGKPDPGRKRRGSPQEEGGLRVSAAAR
		ł	1	ļ	ļ	LLCSGANRCKVLVRQNSTPNTQQPAVHPSTP
	j	I	-		i	PSRPLPQAGRCLVAPLRPHPDWVAAKTLAKA
		ļ	j		!	LRAPGKPWRLAAPSPLGDLGAPGLPGPSTAP
		j	j		į	RTLSVEEPGVECNQLCLYADVTDPVLCLGQK
]	ı	1		1	DPGVEGKHCEKEKISSSKELKHVHAKSEPSKP
		- 1	- 1	l		ARRLSESLHVVDENKNESKIEREHKRRTSTPV
		ļ	- 1			IMEGVQEETDTRDVKRQVERSEICTEEPOKO
					———	

NO: of No: of bod peptide exide sequence peptide lusts in location location corresponding of sequence uence unce unce unce unce unce unc	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence unce dide seq uence unce unc		, ,					D=Aspartic Acid F=Glutamic Acid
Sequence	nucl-	peptide	1				F=Phenylalanine G=Glycine H=Histidine
uence 09496 09496 01st amino acid residue of 09496 09496	cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine.
mino seid reidue of peptide sequence peptide sequence seq	seq-	uence		09/496	correspondi		
residue of pepide pepide pepide pepide pepide pepide sequence RYT-Possible nucleotide decision, "possible nucleotide decision, "possible nucleotide insertion KSTI.KNEKH.IKKDDSETPHI.KSI.LKKEVKSS KEKPEREKTPSDDK.ISVKHKYKGDOMIRKTO DETELHISSEG(ILX VERNICK) GOOTALSD TERKSKHRINERKI.SVLCKGOGRVESVIIKTDE WYRKENNKKERKI.SAEKKAEHKSARSSING IQKDSI.GSKOHGITLORRSESYSEDKCDMDST HORST WINKENNKKERKI.SAEKK.KAEHKSARSSING IQKDSI.GSKOHGITLORRSESYSEDKCDMDST MADDRILK PEPEVHIKERERKI.SHEKSI.CHKS.KAEKK.KS.K.KS.K.TQGKQVK.VVETELQBGATKCATTPKPD KEKNTEEDDBEGKORKS.WDERFFEETQVESTER KS.KS.K.KS.K.KS.K.KS.K.KS.K.KS.K.KS.K.	иелсе	i		914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
peptide sequence Possible nucleotide deletion, \possible						of peptide	T=Threonine, V=Valine, W=Tryptophan,
sequence]]		sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
KSTILKNEKHI KKDDSETPILKSLIKKEVESS RKEPREKITSEOKL SVIERNYKOLOMIKITO DETELHSSEKGI KVEERNIQKOS QOYKLSSDDK TERKSKRINFIREL SVIERNIQKOS QOYKLSSDDK TERKSKRINFIREL SVIERNIGKEVESVIERTDE NVRKENNIKKERRIS SAEKTA EHKSRESDISK (JKOSI GOKQHOTIL ORGENSESYSEDK COMDIST NMDSNIKPEEVVHKERRIT SLLEERICH VLKS RKATOKO VAVVETEL QEGATKQATTEKED KERNITERNOSKORKSK VEDKEPTEET GVEPV LETASSSAHSTOKDSSHRAKI PLAKEKYKSD KOKSTERNISKORKSK VEDKEPTEET GVEPV LETASSSAHSTOKDSSHRAKI PLAKEKYKSD KOKSTER LERKSDEMOKSKSLKHSKSDLIKKE ENKSDDKOKOVDSSHEKARGNSSLMEKKI. SRRLCENRRGSLSQEMKOKEKLA ANTLSTT SOSSLORPKKSGDMTLIPEQEPMEIDSEFGVEV NYFVSKSTONENNISKQDDISSENMOKOKTS ATVOKDELRICTADSKATAPAYKPGRGTOV NSEKHADHSTLIKKMINGSAVSKIMPGE KEPHHIGTTEVNIDSETVIRMIL SAPSENDRY OKNIKINTAGEHVAGOATLEHSTINLDSSPS LSSVTVVELRESVDPDVPLIPDKRIVLEGSTA ATSPADHSALPNOSLIVESEPLLKTSSNEGG EGFTVDTPAKASTISKRRIPPBAHQATLLDGKO GVIMPLOSKLTGVIVENENTKEGGI VDMA KKENDLNAEPHLGOTIKATVENGEKDGI AVD HVUGLINTERYAFTVK IKHRESPEKVIDISD VERNINESVDTSAGSGSAPSVLHORNOQTE DVATOTPBRAKETSTENDENDATISPK AOPATITISSETROSEVALPCTSIEABEGI LIGT HRANNELVAGASECTVFAAAEGGAVVTE GFAESITLTSTKEGGSGCAVASESIDENTAGI LAVHAVKERAVNISVVTEEKDDAVTSAGSEE KCOSLSKINSBEVEGTTITISEVESDOAVTSAG TERRAGISSESEDVSGOGNMMRMOPKKETEG TVTCTGABGRSDNYVLOSVTGAGPREERMYT GAGVULGDNDAPPGTSASQEGGGSVNDGTE GESAVTSTGATEGEGOVIDTOTE GEASTITISTISCUSSORAVISAGSERAND LAVHAVKERAVNISVVTEEKDDAVTSAGSEE KCOSLSKINSBEVEGTTITISEVESDOAVTSAG TERRAGISSEDVSGOGNMMRMOPKKETEG TVTCTGABGRSDNYVLOSVTGAGPREERMYT GAGVULGDNDAPPGTSASQEGGGSVNDGTE GESAVTSTGAKEDEGEGDVIVTSTGREEMIT GAGVULGDNDAPPGTSASQEGGGSVNDGTE GESAVTSTGARGEBEEGDVIVTSTGAREEMIT GAGVULGDNDAPPGTSASQEGGSVNDGTE GESAVTSTGARGEBEEGDVIVTSTGAREEMOT JOESSAKGIVESSVTSAVSGKDEVTPVPGGC GERMTSAASDQSDGGLEAVATSAGSTED UNITERSTGAREEDEGEDVIVTSTGAREEMOT JOESSAKGIVESVTSAVSGKDEVTPVPGGC GERMTSTAARDONSDAGTEKAGSKDT DICSSAKGUPESVTSAVSGKDEVTPVPGGC GERMTSTAARDONSDAGTEKAGSKDT DICSSAKGUPESVTSANSGREDEDUTSWE NEGOGLAMTTASAGDITANDSSAGGREDONATISTS TSECCEAVMIGAVLORDERALTTREVEDLSDA ALISTSTAECHENDALTADREFORLADRENO GDLSATEVSKRIKVPMFSLAERNOCCCPGPV GGREPPOPLANSTEGERONSPHERNSAGGGSST ASVAGGRUE	1						/=possible nucleotide deletion, \=possible
KEKPERKTYSEDKLSVKHYKYGELOMINTO DETELHSSEKGIAVENJOKOSQOTKLSSDDK TERKSKHRINERKLSVLGKOGKPVSEVIRKTDE NYKERNIKERKILSVLGKOGKPVSEVIRKTDE NYKERNIKERKILSVLGKOGKPVSEVIRKTDE NYKERNIKERKILSKEKTKSLLEKEKLVLKS NEMPINKERKILSKEKTKSLLEKEKLVLKS KSKTOGKOVKVVETELOEGATKOATTEKED KEKNTEERNISSEKQEKSKVEDKEPETOVEPV LETASSSAHTOKDSSEKPAKSKVEDKEPETOVEPV LETASSSAHTOKDSSHAKUR JA AKEKYKSD KOSTSTELERKLSOGHEKSKILKLUSKKUKKD ENKSDDWORKEVDSSHRAKUR JA AKEKYKSD KOSTSTELERKLSOGHEKSKILKLUSKKUKKD ENKSDDWORKEVDSSHEKARONSSLMEKKL SRRLCENBRGSLSQEMAKGERKLAANTLSTI SOSSLQRYKASODMILTEPGEPMEIDSETOVEV NYFEVSKTODNENNISGOIDSENMKOKTS ATVOKOELRICTADSSTELERTINLOSSEPS EXPINERTITEVINDISETVIRMLLSAFSENDRY OKNILKNITAAEHVAQGDATLEHSTINLDSSPS LSSTVTVYLRESSYDPVILTOMSTALDSSKOG GEFTVYTVYLRESSYDPVILTOMSTALDSKOG GEFTVYTVYLRESSYDPVILTOMSTALDGKO GKVIMPLOSKLITGVIVENENITKEGGLVDMA KKENDLAMPLKKYTIKATVENKOKKOGIAVD HVVGLNTEKYAETVKLKHRSPGKYKDISID VERRINENSEVDTSAGSASPSVLHORNOOTE DVATGPRRAEKTSVATSTEOKOKOTALSPVK AGPATITTSSETROSEVALPTSIEADGELIGT HSRNNPLHVGAASECTVFAAAEEGGAVVTE GFAESETTLTSTKEGGSECAVASESGELIG EXPRINTISSETTLSSTKOGSECAVASESGERAAD LAVHAVLEANVNSVVTEEKDDAVTSAGSEE KOOSLSRBSSIPSTITTISTEVSSTOGAVTSAG TEIRAGSISSEEVDGSQORMMRMORKETGG TVTCTGAGERSDNPVVLVAGOPCD DGGVTSTGARESGEGAVATSGEG GESATVSTGTGGEGSDATSGESGEGAVS SESEENGESAADSTVAKGGTNVPLVAAGPCD DGGVTSTGAGESGEDVTTSTEGKSCHONLOH ASTCTGLGERSSDNPVLVAKGGNVDLVAGOPCD DGGVTSTGAGESGEGAVTSEGEDDUTSVE VERRENSTGATSGESGEAVASSEGGANS SESEENGESAADSTVAKGGTNVPLVAAGPCD DGGVTSTGAGESGEDVTSTGARSEGGAVTAG TERAGSISSEEVDGSQORMMRMORKETGG TVCTCGAGERSDNPVLYAKGGNNDLVAAGPCD DGGVTSTGAGEEBEEDDUTSVE NECOGLAMATASAGITISSPEGGEAVTSAG TERAGSISSEEVGSAGSGGSGGSGANGAMENGENGE ASTCTGLGERSSDRPVLYAAGPCD DGGTSTGAGEEBEEDDUTSVE NECOGLAMATASAGGEBPVTPVPQGCC GGAVTTSGARSDGAGGGATSTGSBSSEGAMS SESEENGESAANDSTVAKGGTNVPLVAAGPCD DGGTSTGAGEEBEEGDVTTTGAGGGST ANSAGRGLBCAMANSPARABERQCHAMSTST TSGCCAMMIGNALAGGGST ANSAGRGLBCANANSPARABERQCHEGGAST GVANGAGRANDARABERGCTGGGSTAAGGGST ANSAGRGLBCANANSPARABERGCTGGGST ALUNDAGGRANDARABERGCTGGGGST ALUNDAGGRANDARABERGCTGGGTSAGGGST ALUNDAGGAGGAAGABARABARABARABERGCTGGGTSAGGGST					sequence		
DETELHSSEKGLKVERNIQKORGNEVSVETIKTDE TOTRKSKRINERKILSVLGKORGNEVSVETIKTDE NYRKENNIKKERRISASEKTKAEHKSRRSSISSIK (VORSIG GSKQHGHTI-QRSESSYSEDKOMDST NMDSNLKPEEVVHKEKRITKSLLEEKL VILKS SKKTQG (VVKVVETEL DEGATKQATTFEDD KEKNTEENDSEKQEK SKVEDKEFEETOVEPV LETASSSASHTOKONSSHARAL PLAKEKYRSD KOSTSTRLERKLSDCHKSKSLKIBKSLDIKKKD ENKSDDKODGEVDSSHEKARONSSLMEKKL SRRLCENRRGSLSQEMAKGEKLAANTLSTT SOSSILQRYKKSGDMTLIFEQEFMEIDSEFOVE NTEVSKTQDNENNISHQDIDSENMOKGKTS ATVQKDELRICTADSKATAPAYKPGRGTOV NISBEKHADHRSTLTKAMHIGAN SKEMINGKTS ATVQKDELRICTADSKATAPAYKPGRGTOV NISBEKHADHRSTLTKAMHIGAN SKEMINGKTS ATVQKDELRICTADSKATAPAYKPGRGTOV NISBEKHADHRSTLTKAMHIGAN SKEMINGKTS ATVQKDELRICTADSKATAPAYKPGRGTOV NISBEKHADHRSTLTKAMHIGAN SKEMINGKTS ATVQKDELRICTADSKATAPAYKPGRGTOV NISBEKHADHRSTLTKAMHIGAN SKEMINGKTS ATVQKDELRICTADSKATAPAYKPGRGTOV NISBEKHADHRSTLTKAMHIGAN SKEMINGKTS ATVQKDELRICTADSKATAPAYKPGRGTOV NISBEKHADHRSTLTVIBETVIRMILSAPSENDERV QRNLKINTAAEEHVAGOATLEISTINDSSPS LSSVTVVPLRESVDPDVPLFPKRVLGSTA SISPADHRSLPONTAESTEVIRMILSAPSENDERV QRNLKINTAEEHVAGATAGEVARENDERV QRNLKINTAEEHVAGATAGEVARENDERV REPHREMSEVDTAGSGSAPSVLHORNOQTE POPATGRAEKTSVATISTEKDKENDTULSPV AGPATITISSETROSEVALPCTISEADEGLIGT HENDRIPH HYGAEASECTVFAAAEEGGAVTE GFAESETFLTSTIKEGESGCAVAESEDRAADI LAVHAVKIEANVNSVTEEKDAVTSAGSE KCDGSLSRDSEIVEGTTIFISEVSBDAVTSAGSE KCDGSLSRDSEIVEGTTIFISEVSBDAVTSAG TEIRAGISSESEDVGSOORMMRMGRKETEG TVTCTGAEGRSDNYTEGVTGAGFREERMTT GAGVULGINDAPOTSASGEGGSVANDGTE GEASTTITITEGESTORJONDERVENDERDA ADIA LAVHAVKIEANVNSVTEEKDDAVTSAG TEIRAGISSEDVGSOORMMRMGRKETEG TVTCTGAEGRSDNYTEGVTGAGFREERMTT GAGVULGINDAPOTSASGEGGSVANDGTE GEASTTITITEGESTORJONDERVENDERDA ADIA LAVHAVKIEANVNSVTEEKDDAVTSAG TEIRAGISSEDVGSOORMMRMGRKETEG TVTCTGAEGRSDNYTEGVTGAGFREERMTT GAGVULGINDAPOTSASGEGGSVANDGTE BETTLETTSTGATTGAGTGGGSTALARTE ENMEGTIVETRIFARANSSGNDSOLTAS SSEERDGSAMDSTVAKEGTINONSLAGGKENGE VLISTSTTINDTTPOVSATORPEAN SANGGNDSOLTAS RSEENDECAMITSTIGEERDVILTSTAPEGONDOLTAS RSEENDECAMITSTIGEERDVILTSTAPEGONDOLTAS RSEENDECAMITSTIGEERDVILTSTAPEGONDOLTAS RSEENDECAMITST							KSTLKNEKHLKKDDSETPHLKSLLKKEVKSS
TERRISKIRNIFERLISVILGKOGKPVSEVIIKTIDE NYKERNIKKERKILSAEKTKAEHKERSISSIK NOKDSILGSKOJHOTILORRISESYSEDKCDMDST NOMDSILKPEVVHKERKRITKSILLEKILVILKS KSKTOGKOVKVYETELOEGATKOATTPKED KEKNTEENDSKEQKSKSVEDKPETETOVEPV LETASSSAHTOJKDSSHRAKI PLAKEKYKSD KOSTSTRLERKLSDGHKSKSILKHSSKDIKKKD ENKSDDKDGKEVDSSHRAKI PLAKEKYKSD KOSTSTRLERKLSDGHKSKSILKHSSKDIKKKD ENKSDDKDGKEVDSSHRAKI PLAKEKYKSD KOSTSTRLERKLSDGHKSKSILKHSSKDIKKKD ENKSDDKDGKEVDSSHRAKI PLAKEKYKSD KOSTSTRLERKLSDGHKSKSSLAHEKIL SRRICENRKGSISQEHAKGEEKLAANTI.STJ SOSILORPKKSGOMTLIFEGEPWELDSEFOVE NVFEVSKTODNRNNSHQDIDSENMKOKTS ATVOKDELRICTADSKATAPAYKPEGTOV NSNSEKHADHRSTLTKKMHIGSAVSKMNPGE KEPHHRGTTEVNIDSETVIRMILLASPENDRV OKNILKNTAAEHVAQGDATLEHSTNLDSSPS LSSVTVYVPLRESVDPVPIPDKRIVLGSTA STSPADHSALPNOSILTVRESEVILKTSDSKEGG GEFTVYDTPAKASTISKRHPBAHQATILDGKO GKVMPLOSKLTGVVERENILKTSDSKEGG GEFTVYDTPAKASTISKRHPBAHQATILDGKO GKVMPLOSKLTGVVERENILKTSDSKEGG GEFTVTDTAAKASTISKRHPBAHQATILDGKO GKVMPLOSKLTGVVERENILKTSDSKEGG GEFTVTDTAAKASTISKRHPBAHQATILDGKO GKVMPLOSKLTGVVERENILKTSDSKEGG VERNIENSEVDTSAGSGSAPSVLHORNOQTE DVATGFRRAETSVSVATISTEGKORVAVINSID VERRINESPVDTSAGSGSAPSVLHORNOQTE DVATGFRRAETSVSVATISTEGKORVAVINSID VERRINESPVDTSAGSGSAPSVLHORNOQTE GFAESETFLTSTKEGESGCAVAESEDRAADI. LAHHAVKIEANVINSVYTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTTFISEVESDGAVTSAG TERAGSISSEEVDGSOMMMRMGKKETG TVTCTGAGGRSDNFVICSVTGAGPREERMYT GAGVULGNDAPPGTTSAGGEGGVAVTSAG TERAGSISSEEVDGSOMMSMGMKKETG TVTCTGAGGRSDNFVICSVTGAGPREERMYT GAGVULGNDAPPGTSASGEGGSVONDGTE GESAVTSTGTEDGEGBASCTGSEDSSEGFAIS SESEENEGSAMDSTVAKEGTONPUVEH VEAEAGAAMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTFVFGACE GEGAVTSTTITYTTYTGVSATDVSGGLSDALRTE ENMEGTTAAASOSGCARSVILGTTSTGCHEIGH ASTCTOLGESSEEVULGTSTSTGLORDGOTTVEH VEAEAGAAMNANENNVDSMSGTEKGSKDT DICSSAKGIVESEVERADDVITSTGRONEIGH ASTCTOLGESSEEVULGTSTTGTVENDELODA ANTSTTNDYTTGVSATGGTSSTSAVSCCEVTFVFGACE GEGAVTSTTDTYTTGVSATGREDGEGDVTYTTGRENGICH HVAGAGGEBEBELGEBOVTTSTGRONEIGH ASTCTOLGESSEEVULGESBEGGABOSCTALSTSTAGECCAMSS GOVESSEGSAMDSTVAERERMGDTAMISTS TSGCCAAMGGAVLOGDERLITTIKEVEDLDAA ALISTSTABCAPRISASDDRRENOUTADNPEGN GDLASTEVSKHKVPMSSL					•		KEKPEREKTPSEDKLSVKHKYKGDCMHKTG
NVRKENNKKERRISASERISKERSRSDSIS (AVDSIGSGGGITICRRSESSYSEDKOMDST NMDSNLKPEEVVHKEKRITKSILEEKL VILKS KSKTOGKOVKVVETELDGGATKOCHTYRPD KEKNTEENDSEKORKSKVEDKPFEITGVEPV LETASSSAHSTOKDSSIPRAKI PLAKEKYKSD KDSTSTREENLSDGRKSSKLKHSKDIKKN ENKSDDKDGKVDSSIHEKARGNSSLMEKKI. SRRLCENRRGSLSQEWAKGEKLAANTLSTT SGSSLORPKKSGDMTLIPEOEPMEIDSEFGVE NVFEVSKTODNINNSHQDDISSIMKOKIS ATVOKDELRICTADSKATAPAYKPGGTGV VNFEVSKTODNINNSHQDDISSIMKOKIS ATVOKDELRICTADSKATAPAYKPGGTGV VNSEKHADHRSTLKKMHIGGAVSKIMMPGE KEPHRGTTEVHIDSETVIRMLLSAPSENDRY QENLKINTABEHVAGOPATLEISTILDSSPS LSVTVVPLRESYDPDVIPLFDKRTVLEGSTA STSPADHSALPNOSLTVRESEVLKTSDSKEGG GEFTVUTTAKASTISKRHPBAHQATLLDGKQ GKVIMPLOSKLTGVIVENENITKEGGLVDMA KKENDLDARPPLKQTIKATVENGKKOGIAVD HVVGLNTEKYABTVKLKHKRSPGKYKDISID VERNENSEVDTSAGSGASPSVLHORNOOTE DVATGPRRAEKTSVATSTECKDKDVTLSPVK AGPATITTSETRGSEVALPTSIEADEGLIGT HSRNNPLHYGARASECTVFAAAEEGGAVVTE GFAESETTLTSTKEGGSECAVAESEBGLIGT HSRNNPLHYGARASECTVFAAAEEGGAVVTE GFAESETTLTSTKEGGSECAVASEBGRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTTTISEVESDAVTSAGSE KCDGSLSRDSEIVEGTTTTISEVESDAVTSAGSE KCDGSLSRDSEIVEGTTTTISEVESDAVTSAGEE GFAESTTLTSTKTGGSECAVASESBORAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTTTISEVESDAVTSAGGE GEGAVTSTGGTEBGEGFAATSTAGFEREMYT GAGVULGDNDAPPGTSASQEGGGSVNDOTE GESAVTSTGTEBGEGFASCTGSDSSEGFAIS SESEENGESAMDSTVAGFTVSPHACOPCD DEGIVTSTGAKEEDEEGEDVYTSTGRCNEIGH VEARAGAAMMANENNVDSMSGTEKGSKDT DICSSKGRIPSSTVANGSVTTSTEGKEDDLITTSVE NECCDGLMATTASGDTINNSLAGGKNOGK VLINSTSTIND YTTGVASTDPFOGGE GPWTSAASDOSDGLEKVEDTTISTGLVGGS YDVLVSGEVPECEVANTSPEKEDDLITTSVE NECCDGLMATTASGDTINNSLAGGKNOGK VLINSTSTIND YTTGVASTDPFORGE GPWSSAVPGEGDFSVTRAERMGDTAMISTS EGECPAMIGAVLQDFRLTTREFUELDA ALISTSTAECHPISASDRREENGUTADNPEGN GGKEPGPVLAVSTEEGDROFSVTRESAGGGST ASYSAGGRUESNTAALSDATTRAGGSST ASYSAGGRUESNTAALDATARGGSST ASYSAGGRUESNTAALDATARGGSST ASYSAGGRUESNTAALDATARGGSST ASYSAGGRUESNTAALDATARGGSST ASYSAGGRUESNAALBANDARDARDATARGGSST ASYSAGGRUESNAALBANDARDARDARDARDANSPARD	1			ĺ			DETELHSSEKGLKVEENIQKQSQQTKLSSDDK
IOKDSLGSKOHGITLORRSESYSEDKCDMOST NMDSNLKPEEVVHKERRITKSLERKLVLKS KSKTOGKOVKVVETELGEGATKGATTIKVD KEKMTENDESKORKSKVEDELPFETGVEPV LETASSSAHSTOKDSSHRAKLPLAKEKYKSD KOSTSTRLERKLSDGHKSKRLPLAKEKYKSD KOSTSTRLERKLSDGHKSKRLSKSKDIKKKL BIKSDDKDGKEVDSSHERAKNISSKDIKKKL SRRLCENRROSLSGEMAKGEKLAANTLST SOSSLGPKKSGDMTLTPEGPREIDSEPGVE NVFEVSKTODNRNNSHODDSENMKOKTS ATVOKDELRTCTADSKATPANYFORGITOV NSNSEKHADHRSTLTKKMHIQSAVSKMPGE KEPHIRGTTEVNIDSSTVPHKILSSTSA STSPADHSALPNGSLTVRESEVLKTSDSKEG GEFTVDTPAKASTSKRHPPSAHOATLLDGKO GKVMPILGSKLTGVTSPENTKEGGLVDMA KKENDLAGEPILKGTKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHRSPGKVKDISID VERNNENSEVITSAGSGAPSVLHQRNGOTT DVATGPRAEKTSVATSTEGKDKDVTLSFYK AGPATITSETRGSEVALPCTSIEABGGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERNNENSEVITSAGSGAPSVLHQRNGOTT GFASSETILTSTKEGESGCAVAESEDRAADL LAVIHAVKEANVINSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFIEEVESGAVTSAG GESSETVTGGRESGEVALPCTSIEABGGIAVD GFASSETILTSTKEGESGCAVAESEDRAADL LAVIHAVKEANVINSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFIEEVESGAVTSAG GESATSTGTEGGEGGPAVTSTGGGREERMYT GGAVVLGDNDAPGTSAGSGGAPSVLHOOTE GGSAVTSTGTEGGEGGEVATSTGCKORDOTT GGSAVTSTGTEGGEGGEVATSTGCKORDOTT GGSAVTSTGTEGGEGGEVTYSTGGGREERMYT GGGVVLGDNDAPGTSAGSGGASSGASS SESEENGESAMDSTVAKSTTTGLVGGS TVTCTGAEGRSINFVICSVTGAGFREERMYT GGGVVLGDNDAPGTSAGSGGGSTONMENGEKTGG TVTCTGAEGRSINFVICSVTGAGFREERMYT GGSAVTSTGTEGGEGGEVTSTGCRGSSSGGAIS SESEENGESAMDSTVAKSTTTGLVGGS TVTCTGAEGRSINFVICSVTGAGFREERMYT GGSAVTSTGTEGGEGGEVTSTGGRGSSGTAIS SESEENGESAMDSTVAKSTTORVAHOTE GFASATSTGTEDGEGPBASTGGSDSSGGAIS SESEENGESAMDSTAAKSGONDOTTE GGSAVTSTGTEDGEGPBASTGGGSDSCGATS SESEENGESAMDSTAAKSGONDAGTE GGSAVTSTGTEDGEGPBASTGGGSDSCGATS SESEENGESAMDSTAAKSGONDAGTE GGSAVTSTGTEDGEGPBASTGGGSTAAS SESEENGEGAMDSTAASGGNOGGNAMATHENTICLVGGS TVDLVSGEVEGVAATSTSKERDEDGGDSTAAS SESEENGEGAMDSTAASGGRONGAN VLISTSTTNDVTPCGALSDALGGRNOGK VLISTSTTNDVTPCGALSDALGGRNOGK VLISTSTTNDVTPCGALSDALGGRNOGK STAASGGGLGGGANGTAGGERGERMYT STGCCGAAMIGGAV QUEDALTTREAGGRNOGK STAASGGGLGGGANSTAALEGGGTGGGGSST ANSVAGGGLGGGANSTALBEROGTAGGGSST ASSYAGGGLGGGANSTAALBEROGTGGGSST ASSYAGGGLGGGANSTALBEROGTGGGSST ASSYAGGGLGGGANSTA							
NMDSNLKPEEVYKKEKRRTKSLLEEKLVLK KSKTOGKOVKVVETELQEGATKQATTPKPD KERMTEENDAGKORK SKVEDKPFEETGVEPY LETASSAHTOKDSSHRAK JH AAKEYKSD KERSSAHTOKDSSHRAK JH AAKEYKSD KDSTSTLEEKLSDGHKSRSLKHSKDIKKKD ENKSDDKDGKEVDSSHEKARONSSLMEKKL SRRLCENROSIS.GMAKGEKLAANTLSTI SOSSLORPKKSCDMTLPEQEPMEIDSEPGVE NVFEVSKTQONRNNSHQDDSEDMKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTOV NSNSEKHADHRSTLTKKMHIQSAVSKMPGE KEPHRGTTEVVIDSETVHRMLSAPSENDKY KEPHRGTTEVVIDSETVHRMLSAPSENDSPS LSSVTVVPLRESYDPDVPLPDKRTVLGSSTAR SISPADHSALPPQSLTVRESEVLKTSDSKEG GFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSLTGVTERSVLKTSDSKEG GFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVTERSTVLKKKHGGGLVDM HVGLNTEKVAETVKLKKKRSPCKYKDISIB VERRNENSEVDTAGGSGAPSVLHGRNOQTE DVATOPRRAEKTSVATSTEGKDKDVTLSVKY AOPATTTSSETTQSEVALPCTSTEGALDGLIGIT HSRNNPLHVGAEASECTVFAAAEEGGAVDA LAVHAVKRANVSVYTEEKDDAVTSAGSE KCDGSLSRDSELVEGTTTIFSEVSBACTSAG TEIRAGSISSEEVDGSQONMRMMGPKKETEG TTVTCTGAGERSDNFVICSVTGAGRREERWYT GAGVVLGDNBAPFGTSAGSGGSOSPOTOTE GESAVTSTGTEGGEGCAVASCATSAG TEIRAGSISSEEVDGSQONMRMMGPKKETEG TTVTCTGAGESTTHTSTEVSGOSVAPSAG TEIRAGSISSEEVDGSQONMRMMGPKKETEG TTVTCTGAGERSDNFVICSVTGAGRREERWYT GAGVVLGDNBAPFGTSAGSGGSOSVDGTE GESAVTSTGTEGGEGPAASTGSGSSSGGTVAB SESSERGSAMDSTVAEGTTNFVIVAGCPCD DEGIVTSTGAKEEDEGEDVTSTGRGREIGH VERAGAGAMNANENNVDSMSGTEKGSKDI DICSSKAGGVESSVTVAEGTNFVIVAGCPCD DEGIVTSTGAKEEDEGEGDVTSTGRGREIGH VERAGAGAMNANENNVDSMSGTEKGSKDI DICSSKAGGVESSVTVASCKGEVTPVPVGGCE GFMTSAASDOSDGLEKVEIDTTSTELVGG SVDVLVSGEVPECGVAHTSPSEKEDDITTSVE NECCOLMATTASGDITNONSLAGGKNQGK VLLISTSTITDVTPQVSAITDVEGGLSDALTRE EMMEGTEVTTEEFFEAPPSAVSGDSGJCTVABFAPP BASEDLASTSKEEKDECALISTISLEECEASVS VSESDREAGTVANDSTAASGGEPSTYPSGCC GGYMSAAVEGGTPVLSTAADFGGMSSTSSVEDC GGVSSAVPQEGGPSVTPAEMGDTAMISTS TSGCCAAMIGGAVLQDEDLTTTSGLEDGS ASTSKEEKDECAMISTSIGEEFFELDISSTFINC GCKEPGPVLAVSTEEGHROPSVHKPSAQQH PSAVCAEKEERHGICEGFGFGAGGGSSST ASTSAGGGLEGNANSPAHLRGFEGTGGGSSST ASTSAGGGLEGNANSPAHLRGFEGTGGGSSST ASTSAGGGLEGNANSPAHLRGFEGTGGGSSST ASTSAGGGLEGNANSPAHLRGFEGTGGGGSST ASTSAGGGLEGNANSPAHLRGFEGTSGGTAA	(i			i			
KSKTOGKOVK-VETELGEGATKGATTPKPD KERNTEENDBEKORS KSVENDERPETGVEPY LETASSSAHSTOKDSSIRAKI PLAKEKYKSD KDSTSTLLERKI SDGHKSRSLKHSKVIKKD ENKSDDKDGKEVDSSIEK-ARONSSIMEKKI SRRLCENRROSI.SQEMAKGEKLAANTI.STT SOSSLGPRKSGDMTLPREQEMEDISPEGVE NVFEVSKTOONRINNISHODDSENMKQKT. ATVQKDELRTCTADSKATPAYKPGGTOV NSNSEKHADHRSTLTKKMHIQSAVSKMIPGE KEPHRIGTTEVNIDSSTVARMLSSTATSKRSHPGREV VOKNLKINTAAEEHVAQGDATLEHSTNILDSSPS LSSVTVYPLRESYDPDVIPLFDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFYVDTPAKASTISKRHIPPAHQATLLDGKQ (GKVIMPLGSKLTGVIVENENTIKEGGLVDMA KKENDLNAEPILKQTIKATVENGKUGIGAV) HVGGLINTEKVAETVIKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGARSVLHORNOGTE DVATOPRRAEKTSVATSTEGKUKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGLASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGGSGCAVAESDRAADL LAVHAVKIKANNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVPGTITTISEVESDAVTSAGS TEIRAGSISSEVDGSQONMMRMGFKKETEG GESAVTSTGTEGGEOPASCTGSGDSSGGAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDGPASCTGSGDSSGGTAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDGEPASTCTGSGSSGGAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDGEPASCTGSGSSGGAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDGEPASCTGSGUSSGGTAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDGEPASCTGSGUSSGGAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDGEPASCTGSGUSSGGAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDGEPASCTGSGUSSGGAIS SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDGEPASCTGSGUSSGALDAL RICTUGGESEGGVALTSTSTALAGGENGENGH ASTCTQLGEESEGVALICSSAEGDSGIGTVVEH VEABAGAAIMNANENNYDSMSGTEKGSKDT SESEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTAAKSGEDGPASCTGSGUSSGAAR SESENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTAAKSGEDGPASCTGSGUSSGAAR SESENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTAASKGEDGPASCTGSGUSSGGAAG STOVLVSGEVPECEVAHTSPSKEKEDEDITSTSVE NECCOLAMITSSIGEFEPELPISSATTIKCAEE LQPVAAAVEERATGPVLISTAADFEGDMPSPAP BAESPLASTSKEEKDECALISTSIAEECGBSVS OVVESDERAGTTVAEEMGGISTSSVEDC GGYPSSAVPQEGDPSVTTAEEMGDTAMISTS TEGCGCAVMGGAVLQDEDITTTGCLSQGS AUISTAAECMPSALISTSIAECFGBVS GGKEPGPVLAVSTEEGHNGDTSVOTAK BURGKERVALLDSLOKEDGSGGTAA BASTAGGGLEGNANSPAHLRGFEGTGAGGSST ASYSAGRGLEGNANSPA				1			
KERNTEENDSEKORKSIKVEDKPETETGYEV LETARSSAHTOKDSSHRAKE IL-AKELYKSED LENSSAHTOKDSSHRAKE IL-AKELYKSED KDSTSTELERLISDGHKSRSLKHSSKDIKKKD ENKSDDKDGKEVDSSHEKARONSSLMEKKL SRRLCENRGSIS.SQMAKGEEKLAANTLSTI SOSSLQRPKKSQDMALGEEKLAANTLSTI SOSSLQRPKKSQDMALGEEKLAANTLSTI SOSSLQRPKKSQDMALGEEKLAANTLSTI SOSSLQRPKKSQDMALGEEKLAANTLSTI SOSSLQRPKKSQDMALGEEKLAANTLSTI SOSSLQRPKKSQDMALGEEKLAANTLSTI SOSSLQRPKKSQDMALGEEKLAANTLSTI SOSSLQRPKSQDMALGEEKLAANTLSTI SOSSLQRPKSQDMALGEEKLAANTLSTI SOSSLQRPKSQDMALGEEKLAANTLSTI SOSSLQRPKSQDALGEENSTAPAYKPGGGTOV NSMSEKHADHRSTLTKKMHIPGAQAYSMANDROW NSMSEKHADHRSTLTKKMHIPGAQAYSMANDROW NSMSEKHADHRSTLTKKMHIPGAQAYSMANDROW NSMSEKHADHRSTLTKKMHIPGAQAYSMANDROW NSMSEKHADHRSTLTKAMHIPGAQAYSMANDROW NSMSEKHADHRSTLTKAMHIPGAQAYSMANDROW NSMSEKHADHRSTLTKSEVLKTSDSKEGG GEFTYDTPAKASITSKRHIPEAHQATLLDGKQ GEVIMPI,GSKLTGVIVENENTKEGGLVDMA KKENDLNAEPHLAKOTIKATVENGKDGIAVD HVVGLNTEKYAETVIKLKHRSPCKYKDISID VERRIPSSEVDTSAGSGAPSVLHQRNQQTE DVATGPRRAEKTSVATSTEGUKDKUTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIGIT SHRNNPLHVQBASECTVFAAAEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADJ LAVHAVKIEANVISVUTEKEDDAVTSAGSE KCDGSLSRDSEVEGTITIESEVSDQAVTSAG GFERSAGSSEEVDGOQONMMRMQFKKETBQ TEIRAGSISSEEVDGOQONMMRMQFKKETBQ TEIRAGSISSEEVDGOQONMMRMQFKKETBQ TEIRAGSISSEEVDGOQONMMRMQFKKETBQ TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASGEGDGSVTDGTE GESAVTSTGHEDEGEPOSTTGRONEIGH ASTCTGLGEESGGVLICSSAGGDSSGGRAIS SSSEENGESAMDSTVAKEGTNYPLVAAGPCD DEGIVTSTGAKEEDEGEDEVTSTGRONEIGH ASTCTGLGEESGGVLICSSAGGDSGGTVADGTE GESAKGTVESSYTSAVSGKDEVTPYPQGCE GPMTSAASDQSDSQLEKVEDTTTSTGLVGGS YDVLVSGEYPECEVAHTSPSKELGEDITTSVE NEECDGLMATTASGDITNONSLAGGKNOCK VLISTSTNDVTTPQVSATTVEEFGARVS SESEENGESAMDSTVAKEGTNYPLXAAGFGC GGKEPGPLAASTSGEEFERDVSTTADFGFPMPSAPP EAESPLAASTSGEEFERDVSTAADFGFGMPSAPP EAESPLAASTSGEEFERDVSTAADFGFGMPSAPP EAESPLAASTSGEEFERDVTTAEMGDTAMTST SEGCCAVMIGGVLQDERLTTTREEDISDA AIISTTAECMPSASDRHEERNQLTANDFGON GDLSAKGTVAVBERNOLAGGRISTSSVEDC GGKPGPVLAAVSTEEGHNOPSVHKPSAGQGH PSAVCAEKEEKHQKECPFGPFAGRGQKESTI HILNAEENVLLINSLCKERDSFTGTGGGSST ASYSAGRGLEGNANSTAHLRGFEGTSGQTAK SSSKSRYLAAVNTGAKAGAKAA	ł						
LETASSAHSTOKDSHRALIPLAKEKYKSD KNSTSTRILERKISDGKKSKILKHSKOMKKND ENKSDDKDCKEVDSSHEKARGNSSLMEKKIL SRELCENREGIS OSEMAKGEEKLAANTLSTP SGSSLQRPKKSGDMTLIPEGEPMEIDSEPGVE NVFEVSKTQDNRNNSNQDIDSENMKQKTS ATVQKDELRICTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHIQSAVSKMNFGE KEPHRGTTEVNIDSETVHRMILSAPSENDRY QKNILKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVPLRESSYDPDVPIFANTLISAPSENDRY QKNILKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVPLRESSYDPDVPIFANTLISGSSPS LSSVTVPLRESSYDPDVPIFANTLIGGGLOVMA KKENDLNAEPHLKQTIKATVENGKLOGAVD HVQLNTEKYAETVKLEKKRSPGKVKDISID VERNIENSEVDTSAGSGSAPSVLHQRNQQTE DVATGPRAEKTSVATSTEGGLOVMA KKENDLNAEPHLKQTIKATVENGKLOGAVD HVVQLNTEKYAETVKLHKRSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHQRNQQTE DVATGPRAEKTSVATSTEGGLOVMA AGPATTTSSETRQSEVALPCTSIEADEGLIGT HSRNMPHVGAASSECTVSAGSAPSVLHQRNQTE GFAESETILTSTKEGESGCAVAESDRAADL LAVHAVKIEANVNSVYTEEKDDAVTSAGSEE KCDGSLSRDSESVEGTTIFISEVSDGAVTSAG TEIRAGSISSEVDGSQONMMRMOPKKETEG TVTCTGAEGRSDNFVICSVTGAGSEE KCDGSLSRDSEVEGTTIFISEVSDGAVTSAG TEIRAGSISSEVDGSQONMMRMOPKKETEG TVTCTGAEGRSDNFVICSVTGAGSEESGFALS SSEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEGGBAVTSIGGNEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAGAABMNANENNVDSMGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVIPPYGGCE GFMTSAASDQSDQLEKVEDTITISTCUTOGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDITINNSLAGGKNOK VLIJISTSTTNDYTPQVSATIDVEGGLSDALRTE EMMEGTRVTTEEFFAMPSAVSGDDSQLTAS RSEENGEAMDSTVAKEGTNVPLVFAGGC GFMTSAASDQSDQLEKVEDTITISTCUTOGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDITINNSLAGGKNOK VLIJISTSTTNDYTPQVSATIDVEGGLSDALRTE EMMEGTRVTTEEFFAMPSAVSGDDSQLTAS RSEERDECAMISTSTIGEFEEIDFISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP RAESPLASTSKEEKDECALISTSIAEECASVS GVVVSENERAGATVMEEKOGGLISTSSVECC EGPVSSAVPQEEGDFSVTPAEEMGDTAMISTS TSEGCEAVMIGGAVLQDEFRITTTRVDLSDA AIISTSTAECMPISASDRIEEENQLTANDPROG GILSATEVSKRKVPMPSLLABNNCCCPGPVR GGKEPGPVLANSTEEGHNGPSVHKPSAQCGH PSAVCAEKEREKHOKECPEIGFFAGGGSSTL HILMAEEKNVLLINSLQKEDESPTLTTAGGSSET ASYSAGRGLEGNANSFAHLRGPEQTSOQCH PSAVCAEKERECHGECDGTAAN	1 1	1	i		ì		
KDSTSTRLERKLSDGHKSKLKHSSKDIKKKD ENKSDDKDGKEVDSSHEKARGNSSLMEKKL SRLCENBRGSLSQEMAKGEKLAANTLSTT SGSSLQRPKKSGDMTLIPEQEPMEIDBEPGYE NVFEVSKTQDNRNNSNGIDDSENMKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHQSAVSKMNPGE KEPHRGTTEVNIDSETVHRMLLSAPSENDRV QKNLKNTAAEHVAQDATLEHSTNLDSSPS LSSVTVVPLRESVPDYDPLFDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLOSKLTGVIVENENTKEGGLVDMA KKENDLNAEPNIKQTIKATVENGKKDGLAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRINESEVDTSAGSGSAPFUKGTNGATVENGKKDGLAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRINESEVTSAGSGSAPKVLORDISID VERRINESEVTHAGNGTGTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADGAVTSGF HRNNPLHVGAEASECTVFAAAEEGGAVVTL GFAESETILTSTKEGESGECAVAESEDRAADL LAVHAVKEANVNSVYTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEVDFOSQONNMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERNVT GAGVVLGDNDAPPGTSASQEGDSVADGTE GESAVTSTGITEDGEGPASCTOSEDSSEGFAIS SSSEENGESAMDSTVAKEGTNVPLVAAGPCD DEGITYSTGAKEGDEEGEDVTSTGRONEIGH ASTCTGLGESSGVLICESAEGDSQIGTVVEH VEAEGAAMMANENNVDSMSGTEKGSVD DICSSAKGIVESSVTSAVSGKDEVTPYPGGGE GPMTSAASDSSOSQLEKVEDTTSTGLVGGS VULISTSTTNDYTPQVSATDVSGGSDALRTE EMMEGTRVTTEEFFARMFAVSGDDSQLTAS RSEENGEGEGEGEGUTTSTGLVGG VLISTSTTNDYTPQVSATDVSGGLSAALTRCASS LQPVAAAVEERATGPVLISTADFEGMPSAPP EAESPLASTSKEEKDECALISTSLAECECASVS GVVVESENERAGTVMEEKDOGGISTSSVEDC EGPVSSAVPQEEGDPSVTPAERMGISTATIKCABS LQPVAAAVEERATGPVLISTADFEGMPSAPP EAESPLASTSKEEKDEGALISTSLAECECASVS GVVVESENERAGTVMEEKDOGGISTSSVEDC EGPVSSAVPQEEGDPSVTPAERMGISTATIKCABS LQPVAAAVEERATGPVLISTADFEGMPSAPP EAESPLASTSKEEKDECALISTSLAECECASVS GVVVESENERAGTVMEEKDOGGISTSSVEDC EGPVSSAVPQEEGDPSVTPAERMGISTATIKCABS LQPVAAAVEERATGPVLISTADFEGMPSAPP EAESPLASTSKEEKDECALITSTSLAECECASVS GVVVESENERAGTVMEEKDOGGISTSSVEDC EGPVSSAVPQEEGDPSVTPAERMGDTAMISTS TSECCCEAVMIGAVLQDEFRI LTTTRVEDLSDA ALISTSTAECMPISASDIRHEENQLTADNPPQGTVA GKEFGPVLAVSTEEGINGPSVHKPSAQQGH PSAVCAEKEEKHQKECPEIGPAGGGKESTL HLINAEKNVLLNSLQKEDKSPETGTAGGSST							
ENKSDDKDGKEVDSSHEKARGNSSIMERKL SRICCNRRGSLQEMAKGERLAANTLSTT SGSSLQFRKKSGDMTLIPEQEPMEIDSEPGVE NVFEVSRTQCNRRNNSRIQDIDSENMKGKTS ATVQKDELRITADSKATAPAYFORGTOV NSNSEKHADHRSTLTKKMHIQSAVSKMRGE KEPHRGTTEVNIDSETVHRMHQSAVSKMRGE KEPHRGTTEVNIDSETVHRMHQSAVSKMRGE KEPHRGTTEVNIDSETVHRMHQSAVSKMRGE KEPHRGTTEVNIDSETVHRMHQSAVSKMRGE KEPHRGTTEVNIDSETVHRMHQSAVSKMRGE KEPHRGTTEVNIDSETVHRMHQSAVSKMRGE KEPHRGTTEVNIDSETVHRMHQSAVSKMRGE KEPHRGTTEVNIDSETVHRMHQSAVSKMRGE KEPHRGTTEVNIDSETVHRMHQSAVSKMRGE KEPHRGTTEVNIDSETVHREVELLGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVMPLGSKLTGVIVENENTIKEGGLVDMA KKENDLNAEPNIKGTIKATVENGKKDGLAVD HVVGLNTEKYASTVALKHKSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHQRNQGTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSSEVALPCTSIEADEGLIGT HSRNNPHVGAEASECTVAAAEEGGAVVTE GFAESETHLTSTKEGESGECVAAAEEGGAVVTE GFAESETHLTSTKEGESGECVAAAEEGGAVVTE GFAESETHLTSTKEGESGECVAAAEEGGAVVTE GFAESETHLTSTKEGESGECVAAAEEGGAVVTE GFAESETHLTSTKEGESGECVAAAEEGGAVVTE GFAESETHLTSTKEGESGECVAAAEEGGAVTE GFAESETHLTSTKEGESGECVAAAEEGGAVTTE GFAESETHLTSTKEGESGECVAAAEEGGAVTE GFAESETHLTSTKEGESGECVAAAEEGGAVTE GFAESETHLTSTKEGESGECVAAAEEGGAVTE GFAESETHLTSTKEGESGECVAAAEEGGAVTE GFAESETHLTSTKEGESGECVAAAEEGGAVTE GFAESETHLTSTKEGESGECVAAAEEGGAVTE GFAESETHLTSTKEGESGECVAAAEEGGAVTE GFAESETHLTSTKEGESGECVAAAEEGGAVTE GFAESETHLTSTKEGESGECVAAAEEGGAVTE GFAESETHLTSTKEGESGECVAAAEEGGAVTE GFAESETHLTSTKGEGSGECVAAAEEGGAVTE GFAESETHLTSTGEGEGETHOTSTVLVAAGPCD DEGIVTSTGGEGEGASCTGSGDSSEGRAIS SESEENGESAMDSTVAKEGTHVPLVAAGPCD DEGIVTSTGGEGEGEDVYTSTGGROEGCE GFMTSTGGTEDGEGFASCTGSGDSSEGRAIS SESEENGESAMDSTVAKEGTHVPLVAAGPCD DEGIVTSTGGAPEGEROVTSTGGROEGCE GFMTSAASDQSDSQLEKVEDTITSTGLVGGC GFMTSAASDQSDSQLEKVEDTITSTGLVGGC GFMTSAASDQSDSQLEKVEDTITSTGLVGGC GFMTSAASDQSDSQLEKVEDTITSTGLVGGC GFMTSAASDQSDSQLEKVEDTITSTGLVGGC GFMTSAASDQSDSQLEKVEDTITSTGLVGGC GFMTSAASDQSDSQLEKVEDTITSTGLVGGC GFMTSAASDQSDSQLEKVEDTITSTGLTVGC NCACCACCACCACCACCACCACCACCACCACCACCACCAC	1				l		KDSTSTRLERKI SDGHKSKSI KHSSKDIKKKD
SRRLCENRIGGLSQEMAKGEEKLAANTI.STP SGSSLQRFKKSGIMTI.EPGPEIDSPEYGVE NVFEVSKTQDNRNNNSHQDIDSENMKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHQSAVSMKMPGE KEPHRGTTEVNIDSETVHRMLI.SAPSENDRV QKNILKNTAAEHVAQGDATULEHSTNLDSSPS LSSVTVVPLRESVDPDVPLFDKRTVLEGSTA STSPADHSAL.PNOSLTVRESPLK.KTDSKEGG EGFTVDTPAKASITSKRHIPAHQATLLDGKQ GKVMPH_OSKLTGVIVENENTKEGGLVDMA KKENDLNAEPHLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRINENSEVDTSAGSGSAPFKVHGNNGCTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETROSSVALPCTSLEADEGLLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETTLTSTKEGGSECAVAESEDRAADL LAVHAVKEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEEVDGSQONMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSAGSGOSVADGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SSESENGSAMDSTVAKEGTNVPL.VAAGPCD DEGIVTSTGAKEEDEEGGDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMMANENNVDSMSGTEKGNSIDT DICSSAKGIVESSVTSAVGKDEVTPVPGGCE GFMSAASDQSSOCJEKVESTTSLAVGROCK VLIISTSTTNDYTPQVSAITDVEGGSDALRTE EMMEGTRVTTEEFFARMSAVSGDBSQLTAS RSEEKDECAMSTISGFEEGDSTSSTELVEG VLIISTSTTNDYTPQVSAITDVEGGSDALRTE EMMEGTRVTTEEFFARMSAVSGDBSQLTAS RSEEKDECAMSTISGFEEGERISSSTISAVGKOGK VLIISTSTTNDYTPQVSAITDVEGGSDALRTE EMMEGTRVTTEEFFARMSAVSGDBSQLTAS RSEEKDECAMSTISGFEEFEISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGMPSAPP EAESPLASTSKERELDECALISTSCATLT EMMEGTRVTTEEFFARMSAVSGDBSQLTAS RSEEKDECAMSTISGFEEFEISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGMPSAPP EAESPLASTSKERELDECALISTSSATTIKCAES LQPVAAAVEERATGPVLISTADFEGMPSAPP EAESPLASTSKERELDECALISTSSATTIKCAES LQPVAAVEERATGPVLISTADFEGMPSAPP EAESPLASTSKERELDECALISTSCATLT EMMEGTRVTTEEFFARMSAVSGDBSQLTAS RSEEKDECAMSTISGFEEFEISSATTIKCAES LQPVAAVEERELDECALISTSORDE GGKPGPGVLAVTSGEGIDRSVHRPSAQGOH PSAVCAEKEERHQKACPEIGPFAGRQKESTI. HLINAEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRILEGNANSFAHLRGPEGTSGOTAK DSVSSIRYLAAVITGAKADDMPPVQGTVA							
SGSSLQRPKKSGDMTLIPEGEMERDSEPGVE NVFEVSKTQDNRNNNSHQDIDSENMKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTGV NSNSEKHADHRSTLTKKMHIGSAYSKMPGE KEPIHRGTTEVNIDSETYHMLLSAPSENDRV QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYPDVPLPDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENTIKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVVGLNTEKY AETVKLKHRSPGVKDISID VERNENSEVDTSAGSGSAPSVLHQRNQQTE DVATGPRRAKITSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETTLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSE KCDGSLSRDSSLVEGTITTISEVESDGAVTSAG TERAGGISSEEVDGSQGNMMRMOPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVULGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSEGPAIS SESENGESAMDSTVAKRGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRQNEIGH ASTCTGLGEESGGVLICSEGGAGSGNDTOHE ASTCTGLGEESGGVLICSEGGAGGSGNDTOHE ASTCTGLGEESGGVLICSEGGAGGSGNDTOHE VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTYPGGCE GFMTSAASDQSSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDDITISTGL TOLGSSAKGIVESSVTSAVSGKDEVTYPGGCE GFMTSAASDQSSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDDITSVE NEECDGLMATTASGDTINNSLAGGKNOGK VLISTSTTNDVTPQVSAITDVFBGGLSDALTER ENMEGTRVTTEEFEAPMPSAVSGGLSDALTER ENMEGTRVTTEEFEAPMPSAVSGLOSQLTAS RSEEKDECAMISTSIGEFEILPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC GGYPSSAVPQEEGDPSVTTAEMGDRCRCPGPVR GGKEPGPVLAVSTEEGFINGSPSVIRLSDAA AIISTSTAECMPISASIDRHEEROLTAMISTS TSECCEAVMIGAVLQEERITTITRVEDLSDA AIISTSTAECMPISASIDRHEEROLTAMISTS TSECCEAVMIGAVLQEERITTITRVEDLSDA AIISTSTAECMPISASIDRHEEROLTAMISTS TSECCEAVMIGAVLQEERITTITRVEDLSDA AIISTSTAECMPISASIDRHEEROLTAMISTS TSECCEAVMIGAVLQEERITTITRVEDLSDA AIISTSTAECMPISASIDRHEEROLTAMISTS TSECCEAVMIGAVLQEERITTITRVEDLSDA AIISTSTAECMPISASIDRHEEROLTAMISTS TSECCEAVMIGAVLAUSTARANDPARGRCPCPPVR GGKEPGPVLAVSTEEGFINGSPSVIRSAGGGNESTIL HLINAEENNYLLNSLQKEDKSPSTGAGGST							SRRLCENRRGSLSOEMAKGEEKLAANTI.STP
NVFEVSKTQDNRNNSHQDIDSENMKQKTS ATVQKDELRTCTADSKATAPAYKPGRGTOV NSNSEKHADHRSTLTKKMHIQSAVSKMNPGE KEPIHRGTTEVNIDSETVHRMLSAJESNDRV QKNLKNTAAEHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYPDDVPLFDKRTVLBGSTA STSPADHSALPNQSLTVRESVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNABEPULKQTIKKATVENGKKDGJAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRINENSEVDTSAGGSAPSVLHQRNQQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGFATITTSSETRQSEVALFCTSIEADGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRADL LAVHAVKIEANNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGISSEEVDSOSQONMMRNGPKETEG TVTCTGAEGRSDNFVICSVTGAGPREERNVT GAGVVLGDNDAPPGTSASQEGJGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTINVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGROEIDH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTERGSKDT DICSSAKGIVESSVTSAVSGEDSVTPVPGGGE GPMTSAASDQSDSLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITISVE NECCOLLMATTASGDITNQNSLAGGKNQGK VLISTSTTNDVTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEFFEAPMSAGDDSDALTAS RSEEKDECAMISTSIGEFELPISSATTIKCAES LQPVAAAVEERATAGFVLISTAJPEGDPAGRPSAP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVESENERAGTVMEEKDGSGITSSVEDC EGPYSSAVPQEEGDPATTAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRITTTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GGKEPGPVLAVSTEEGRINGPSVHKPSAGQGH PSAAVCAEKEERHGKECPEGIPFAGRGKOESTI. HLINAEEKNVLLNSLQKEDKSPSTGATGSGTTAK DSSVSSIKTLAANNSPAHLRGPEGTSGQTTAK DSSSVSSIKTLAANNSPAHLRGPEGTSGQTTAK DSSSVSSIKTLAANNSPAHLRGPEGTTSGQTTAK DSSSVSSIKTLAANNSPAHLRGPEGTTSGQTTAK DSSSVSSIKTLAANNSPAHLRGPEGTTSGQTTAK DSSSVSSIKTLAANNSPAHLRGPEGTTAGGSST							SGSSLQRPKKSGDMTLIPEOEPMEIDSEPGVE
ATVOKOELRTCTADSSATAPAYKPGEGTOV NSNSEKHADHRSTITKKMHIOSAVSKMPRGE KEPIHRGTTEVNIDSETVHRMLISAPSENDRV QKNI.KNTAAEEHVAQGDATLEHSTILDSSPS LSSVTVVPLERSYPDVIPUFDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTIDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLIDGKQ GKVIMPLGSKLTGVIVENENTIKEGGLVDMA KKENDLNAEPNILKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLIKKRSPGKVKDISID VERRIENSEVDTSAGSGSAPSVLHQRNQQTE DVATGPRRAEKTSVATSTEGKDKTULSPVK AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNIPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGVALAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEEVDGSQONMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMYT GAGVVLGDNDAPPGTSAGCGGGSVNDGTE GESAVTSTGITEDGEGPASCTOSEDSSEGFAIS SSBEENGESAMDSTVAKEGTINVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRGNEIGH ASTCTIGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTIS VGGS YDVLVSGEVPECEVAHTSPEKERDEDIITSVE NEECDGLMATTASGDTINQNSLAGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSALRTE ENNEGTRVTTEEFFAPPRSAVSGDDSQLTAS RSEEKDECAMISTSIGEFFELPISSATIKCAES LQPVAAAVERATGPVISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPYSSAVPQEEGDPATTAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRITTIRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADPEGN GGKEPGPVLAVSTEEGRINGPSVHKRSSAQCH PSAVCAKEEEGHRINGFSPTAREMGDTAMISTS TSEGCEAVMIGAVLQDEDRITTIRVELSDA AIISTSTAECMPISASIDRHEENQLTADPEGN GGKEPGPVLAVSTEEGRINGPSVHKRSPSAQCH PSAVCAKEEEGHRINGFSPTAGAGSST ASYSAGRGLEGNANSPAHLROPFGOTSGCTAK DSSSVSSIRVLAANNSPAHLROPFGOTSGCTAK DSSSSSIRVLAANNSPAHLROPFGOTSGCTAK	1	}	1	j			NVFEVSKTQDNRNNNSHQDIDSENMKQKTS
KEPIHRGTTEVNIDSETVHRMLLSAPSENDRV QKNI.KNTAAEEHVAQDATLEHSTNILDSSPS LSSVTVVPLRESYDPDVIPJFDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLIKTSDSKEGG EGFTVDTPJRAKSITISKHPEAHQATILDGKQ GKVIMPLGSKLTGVIVENENTIKEGGLVDMA KKENDLNAEPNILKQTIKKTVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNGQTE DVATOPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEGRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCOGSLSROSEIVEGTITISEVESDGAVTSAG TEIRAGSISSEEVDGSQONMARMGRKETEG TVTCTGAEGRSDNFVICSVTGAGRPEERMVT GAGVVLGNDAPPGTSASQEEGDSVDOTTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGSOGIGTVVEH VEAAEGAAMNANENNDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTSTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITISVE NEECDGLMATTASGDTNONSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALTISVE NEECDGLMATTASGDTNONSLAGGKNQGK VLISTSTTNDYTPQVSAITDVEGGLSDALTE ENMEGTRVTTEFFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFFELPISSATTIKCAES LQPVAAAVERATGPVITAEFREMGDTAMISTS TSEGCEAVMIGAVLQDERLITTEVEDLSDA ALISTSTAECMPISASITISTAEFCEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGFVSSAVPGEEGDFSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDERLITTEVEDLSDA ALISTSTAECMPISASITHEKCAES LQPVAAVERATGPVITAEEMGDTAMISTS TSEGCEAVMIGAVLQDERLITTEVEDLSDA ALISTSTAECMPISASITHEKCAES LQPVAAVERENTOPSVTRAEEMGDTAMISTS TSEGCEAVMIGAVLQDERLITTEVEDLSDA ALISTSTAECMPISASITHEKCAES GREPOPVLAVSTEEGHNOPSVHKPSAQQGH PSAVCAEKEKHKKPMPSLIABNNCCCPGPVR GGKEPGPVLAVSTEEGHNOPSVHKPSAQQGH PSAVCAEKEKHKGEGFFFGAGGGGKESTL HLINAEEKNVILINSLOKEDKSPETGTAAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSQCTAK DSSVSSIRYLAAVNITGALKADDMPPVQGTVA					!		
OKNILKNTAAEEHVAQGDATLEHSTNLDSSPS LSSVTVVPLRESYDPDVIPLPKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPPAHQATLLDGKQ GKVIMPLGSKLTGVIVENENTIKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKIKRSPGKVKDISIDI UVERNIENSEVDTSAGSAPSVLHQRNGQTE DVATOPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKEANVNSVYEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTIFISEVESDGAVTSAG TEIRAGSISSEVDGSQONMMRMGPKKETEG TVTCTGAEGRSDNPVCTVGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRQNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSKDEVTPVPGGCE GPMTSAASDQSBOGLEKVEDTTSTGLYGGS YDVLVSGEVPECEVAHTSPSEKEDEDITISTELYGGS YDVLVSGEVPECEVAHTSPSEKEDEDITISTELYGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDITNONISLAGGRNQGK KULISTSTINDYTPQVSAADSODSOLTAS RSEKDECAMISTSIGEEFELPISSATTIKCAES LQFVAAAVERRATGPVLSTADFEGPMPSAPP EAESPLASTSKEEKDECALISTISLAECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDERLITTITEVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNOPSVHKPSAQQGH PSAVCAKEKEKHGECGFFFAGRGGKESTL HLINAEEKNVLLINSLOKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNITGALKADDMPPVQGTVA			1				
LSSVTVVPLRESYDPDVPLEDKRTVLEGSTA STSPADHSALPNQSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRHIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENTIKEGGLVDMA KKENDLNAEPHLKQTKATVENGKKGGAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAFSVLHQRNQGTE DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAASEGTVAAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEBRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTISEVESDGAVTSAG TEIRAGSISSEVDGSQONMMEMGPKKETEG TVTCTGAEGRSDNFVICSVTIESEVESDGAVTSAG TEIRAGSISSESTOGSQONMMEMGPKKETEG TVTCTGAEGRSDNFVICSVTIESEVESDGAVTSAG TEIRAGSISSESTOGSGONFVVENTOTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESSGVLICESAEGDSGIGTVVEH VEAAEGGAAINANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDTINOPSAGTREGSSTINVE NEECDGLMATTASGDTINOPTOGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFELPISSATTIKCAES LQPVAAAVERATGPVLISTADFEGRMPSAPP EAESPLASTSKEEKDECALISTSIAFECEASVS GVVVESENERAGTVMEEKDGSISTSSVEDC EGFVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDERLITTIKEPDLSDA AIISTSTAECMPISASJDSRIEENOLTADPEGON GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNOPSVHKPSAQQGH PSAVCAKEKEKHKECFGFFAGRGGKESTL HLINAEEKNVILINSLOKEDKSPETGTAAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTTAK DSSVSSIRYLAAVNIGAKADDMPPVQGTVA		[ĺ		ĺ	[
STSPADHSALPNOSLTVRESEVLKTSDSKEGG EGFTVDTPAKASITSKRIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPNLKQTIKATVENGKKDGIAVD HVGCLINTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNQGTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITTSSETROSEVALPCTISEADGGLIGT HSRNNPLHYGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGCAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEEVDGSQOMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEVVTSTGRGNEIGH ASTCTGLGESEGGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVOGCCE GPMTSAASDQSDSQLEKVEDTITSTEGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASQDITNQNSLAGGKNQGK VLISISTSTNDYTTQVSAITDVEGGLSDALRTE ENMEGTRVTTEFFRAPMFSAVSGDDSQLTAS RSEEKDECAMISTSIGEFFELPISSATTIKCAES LQPVAAAVEERATGPVLISTAGEFFELPISSATTIKCAES LQPVAAAVEERATGPVLISTAGEFFELPISSATTIKCAES GVVESENERAGTVUESTIGEFFELPISSATTIKCAES LQPVAAAVEERATGPVLISTAGEFFELPISSATTIKCAES SEENGESHASTISKEEKDECALISTSIAEECEASVS GVVESENERAGTVERECHOSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRILTTIRVEDLSDA AIISTSTTAGENEEROGGTITTRVEDLSDA AIISTSTAECHPISAGDRHEENQLTADNPEGN GDLSATEVSKIRKVPMPSLLABNNCRCPGPVR GGKEPGPVLANSTEEGHNGPSVHKPSAQGGH PSAVCAEKEKHIGKECPEIGPFAGRGGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPQGTVA							QKNLKNTAAEEHVAQGDATLEHSTNLDSSPS
EGFTVDTPAKASITSKRIIPEAHQATLLDGKQ GKVIMPLGSKLTGVIVENENITKEGGLVDMA KKENDLNAEPILKQTIKATVENGKKDGIAVD HVVGLNTERYAETVKLKHKRSPGKVKDISID VERRHENSEVDTSAGSGAPSVLHQRNQQTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEGGAVVTE GFAESETFLTSTKEGESGCAVAESEDRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTTIFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGFREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEGEDVVTSTGGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NECCDGLMATTASGDITNQNSLAGGKNQGK VLISTSTTNDYTPQVSATDVEGGLSDALRTE ENMEGTRVTTEEFFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFFELPISSATITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECLISTSIAECECASVS GVVVESENERAGTVMEEKDDGSGISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLJAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKKEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDSSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSIRYLAAVNTGAIKADDMPPQGTVA	1 1	İ	- 1	1	ļ	ĺ	LSSVTVVPLRESYDPDVIPLFDKRTVLEGSTA
GKVIMPLGSKLTGVIVENENTKEGGLVDMA KKENDLNAEPILKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVALKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAFSVLHQRNGQTE DVATGPRAEKTSATSEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEBRAADL LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREEMMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTIGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGPVTSTGRGNEIGH ASTCTGLGESEGGVLICESAEGDSQIGTVVEH VEAEAGAAMMANENNVDSMGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTITISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDHITSVE NECEDILMATTASDITNONSLAGGRNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFAPMPSAVSGDDSQLTAS RSEKKDECAMISTSIGEFFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIGEFEELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSISLAEECEASVS GVVVESENERAGTVEEKDOSGGITSSVEDC EGPVSSA VPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTTIRVEDLSDA AIISTSTAECMPISASDRHEENQLTADNPEGN GDLSATEVSKIKVPMPSLAENNCCCPOPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAQGH PSAVCAEKEKHGKECPEIGPFAGRGQKESTL HLINABEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLERDERPQCTSQQTAK DSSVSSIRYLAAVNITGAIKADDMPPVQGTVA					1		STSPADHSALPNQSLTVRESEVLKTSDSKEGG
KKENDLNAEPNILKQTIKATVENGKKDGIAVD HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSGSAPSVLHQRNQGTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITTISEVESDGAVTSAG TEIRAGSISSEEVDGSQONMRMGMPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTTAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA ALISTSTAECMPIESAIDTHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCFGPVR GGKEPGPVLAVSKPRSAGOGH PSAVCAEKEEKHIGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEBKSPETGTAGGSTA ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNIGAIKADDMPPVQGTVA		l	i	i	ì	1	GKAMMI GSKI TGAMENENITYEGGI ADAKA
HVVGLNTEKYAETVKLKHKRSPGKVKDISID VERRNENSEVDTSAGSSAPSVLHQRNQGTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESSETFLTSTKEGESGECAVAESDRAADL LAVHAVKIEANVNSVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITITISEVESDGAVTSAG TEIRAGSISSEEVDGSQONMMRMOPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGTIEGEPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEEGDVTSTGRONEIGH ASTCTGLGEESGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSYTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE EMMEGTRVTTEEFFAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFFLPISSATTIKCAES LQPVAAAVERATGPVLISTADFEGMPSAPP EAESPLASTSKEELDEGALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGGEAVMIGAVLQDEDRLITTIVEVDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHRPSAQGGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNIGAIKADDMPPVQGTVA		1	l l				
VERRNENSEVDTSAGSGSAPSVLHQRNQGTE DVATGPRRAEKTSVATSTEGKDKDVTLSPVK AGPATITSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANNNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQOMMRMGMGKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRONEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPYPOGGE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDITTSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTINDYTPQVSATTDVEGGLSDALRIE EMMEGTRVTTEFFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEKDECALISTISIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGGCAVMIGAVLQDEDRLITTRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLJAENNCRCPGPVR GGKEPGPVLASTEGEHNOPSVHKPSAQQGH PSAVCAEKEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSQQTAK DSSVSSIRYLAANTGAHLRADDMPPVQGTVA	[]	i	ŀ	ł	• 1		
DVATGPRAEKTSVATSTEGKDKDVTLSPVK AGPATITSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANNSVYTEKDDAVTSAGSEE KCDGSLSRDSVEGTITFIESEVESDGAVTSAGSEE KCDGSLSRDSVEGTITFIESEVESDGAVTSAGSEE KCDGSLSRDSVEGTITFIESEVESDGAVTSAGSEE TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SSEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEGDEVTSTGRONEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSKDEVTYPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA ALISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKIKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNOPSVIKPSAQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSSFETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSQQTAK DSSVSSIRYLAAVNTGALKADDMPPVQGTVA	li	ľ	i	1	!		
AGPATTTSSETRQSEVALPCTSIEADEGLIIGT HSRNNPLHVGAEASECTVFAAAEEGGAVVTE GFAESETHITSEGGGSCAVAESEDRAADL LAVHAVKIEANVNSVVTEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQGNMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRONEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTSGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEKDGSGIISTSSVEDC EOPPVSAVPQEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNOPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSQQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		l	l	İ	. '		
HSRNNPLHYGAEASECTVFAAAEEGGAVVTE GFAESETFLTSTKEGESGECAVAESEDRAADL LAVHAVKIEANVNSVVTEKDDAVTSAGSEE KCDGSLSRDSEIVEGTITFISEVESDGAVTSAG TEIRAGSISSEEVDGSQONMMRMGPKKETEG TVTCTGAEGRSDNPVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPYPGGCE GPMTSAASDQSDSQLEKVEDTITSTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTINDYTPQVSAITDVEGGLSDALRITE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEFELPISSATTIKCAES LQPVAAAVERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAREMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA ALISTSTAECMPISASDRHEENQLTADNPEGN GDLSATEVSKKHSVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPSVHKPSAGGGT	i i	1	ł	i	•		
LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE KCDGSLSRDSEIVGTITTISEVESDGAVTSAG TEIRAGSISSEEVDGSQGMMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGFREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQOK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGFVLISTADFEGPMPSAPP EAESPLASTSEKEDECALISTSIAEECASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAQQCH PSAVCAEKEEKHGKECPEIGPFAGRQKESTL HLINAEEKNVLINSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		1					
KCDGSLSRDSEIVEGTTTFISEVESDGAVTSAG TEIRAGSISSEEVDGSQOMMMRMOPKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECAISTSIAEECEASVS GVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GGLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHIGKECPEIGFFAGRGSKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGMANSPAHLRGPEGTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		ļ	1	-	i		
TEIRAGSISSEEVDGSQGNMRMGPKKETEG TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAAVEERATGPVLISTADFEGPMPSAPP EASSPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GGLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGFFAGRGGKESTL HLINAEEKRIVLINSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		j	ļ		1		LAVHAVKIEANVNSVVTEEKDDAVTSAGSEE
TVTCTGAEGRSDNFVICSVTGAGPREERMVT GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGGREIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDHITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGISTSSVEDC EGPVSSAVQEEDDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		1]	1	1		
GAGVVLGDNDAPPGTSASQEGDGSVNDGTE GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRONEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDITTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGOKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	ŀ		i				
GESAVTSTGITEDGEGPASCTGSEDSSEGFAIS SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVYTSTGRGEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFFAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGGTTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	l J	1	J		j		
SESEENGESAMDSTVAKEGTNVPLVAAGPCD DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA ALISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA			1				
DEGIVTSTGAKEEDEEGEDVVTSTGRGNEIGH ASTCTGLGESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDQLEKVEDITISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		}	- 1			ì	
ASTCTGLGEESEGVLICESAEGDSQIGTVVEH VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA			- 1	1		ļ	
VEAEAGAAIMNANENNVDSMSGTEKGSKDT DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIK CAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEBCEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA			İ	- 1	.	İ	ASTCTGLGEESEGVLICESAEGDSOIGTVVFH
DICSSAKGIVESSVTSAVSGKDEVTPVPGGCE GPMTSAASDQSDSQLEKVEDITISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		İ	- 1	1	i	- 1	VEAEAGAAIMNANENNVDSMSGTEKGSKDT
GPMTSAASDQSDSQLEKVEDTTISTGLVGGS YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVITEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		j	1		•		
YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE NEECDGLMATTASGDITNQNSLAGGKNQGK VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		1	- 1	i		i	
VLIISTSTTNDYTPQVSAITDVEGGLSDALRTE ENMEGTRVTTEEFEAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSAITIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLLAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		ĺ	. [1	.]	ĺ	YDVLVSGEVPECEVAHTSPSEKEDEDIITSVE
ENMEGTRVTTEEFAPMPSAVSGDDSQLTAS RSEEKDECAMISTSIGEEFELPISSATTIK CAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		į	1		- 1	1	
RSEEKDECAMISTSIGEEFELPISSATTIKCAES LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		[- 1	1		l	
LQPVAAAVEERATGPVLISTADFEGPMPSAPP EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		1	}	ŀ		[
EAESPLASTSKEEKDECALISTSIAEECEASVS GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		- 1	- 1	ł	- 1	Į	
GVVVESENERAGTVMEEKDGSGIISTSSVEDC EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		i	- [1	1	l	
EGPVSSAVPQEEGDPSVTPAEEMGDTAMISTS TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		- 1	ł	1	į.	ļ	
TSEGCEAVMIGAVLQDEDRLTITRVEDLSDA AIISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		\ 	- 1	- 1		.	
ALISTSTAECMPISASIDRHEENQLTADNPEGN GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		1	- 1	- 1	1	J	
GDLSATEVSKHKVPMPSLIAENNCRCPGPVR GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		1			1	į	
GGKEPGPVLAVSTEEGHNGPSVHKPSAGQGH PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		1	- }	j	1		
PSAVCAEKEEKHGKECPEIGPFAGRGQKESTL HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA					j		
HLINAEEKNVLLNSLQKEDKSPETGTAGGSST ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA			- 1		İ		
ASYSAGRGLEGNANSPAHLRGPEQTSGQTAK DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA	}	ĺ	- [- 1	- 1	ľ	
DSSVSSIRYLAAVNTGAIKADDMPPVQGTVA		[1	ſ	
EUCEI DA COO COCENTII VITOTTO DE CONTANT		į	1	}			
Enservative Enservation Enservative Enservation Enservative Enservation Enservative Enservation Enservative Enservation Enservative Enservation Enservative Enservative Enservation Enservative Ense			L				EHSFLPAEQQGSEDNLKTSTTKCITGQESKIAP

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion SHTMIPPATYSVALLAPKCEQDLTIKNDYSGK WTDQASAEKTGDDNSTRKSFPEEGDIMVTVS SEENVCDIGNEESPLNVLGGLKLKANLKMEA YVPSEEEKNGEILAPPESLCGGKPSGIAELQRE PLLVNESLNVENSGFRTNEEIHSESYNKGEISS GRKDNAEAISGHSVEADPKEVEEEERHMPKR
1302	2652	A	10167	321	842	KRKQHYLSSEDEPDDNPDVLDSRIETAQRQC PETEPHATKEENSRDLEELPKTSSETNSTTSRV MEEKDEYSSSETTGEKPEQNDDDTIKSQE EPSLFPFLRPSPARPPPRPPAPFPSPELAGPEPH
						FVFYFFLSYVHPPKELAKYEYMEEQVILTEKG NSTVAGRGTSVRCLSPSPRPLPPLLPLLADLLE DGFGEHPFYHCLVAEVPKEHWTPEGNPSPFP EARETKCYVRSSVGCVEPLTTQAEVTENLDR KNSQOVFKLLKKK
1303	2653	A	10171	206	429	NMILLKKRRILINSLGEGTINGLLDELLETNV LSQEDTEIVKCENVTVIDKARDLLDSVIRKGA RACEICITYI
1304	2654	A	10184	970	1524	LCTLSPGISGTAGSCLTTEPGTELGTSFAQNGF YHEAVVLFTQALKLNPQDHRLFGNRSFCHER LGQPAWALADAQVALTLRPGWPRGLFRLGK ALMGLQRFREAAAVFQETLRGGSQPDAAREL RSCLLHLTLQGQRGGICAPPLSPGALQPLPHA ELAPSGLPSLRCPRSTALRSPGLSPLLH
1305	2655	A	10194	2	394	TDLLGRRFRVDGAAMAACEGRRSGALGSSQ SDFLTPPVGGAPWAVATTVVMYPPPPPPHR DFISVTLSFGESYDNSKSWRRRSCWRKWKQL SRLQRNMILFLLAFLLFCGLLFYINLADHWKG IRNTCT
1306	2656	A	10195	1	410	IPGSTISLEGPLSKWTNVMKGWQYRWFVLDY NAGLLSYYTSKDKMMRGSRRGCVRLRGAVI GIDDEDDSTFTITVDQKTFHFQARDADEREK WIHALEETILRHTLQLQVRVFTWFPDSSLVGA FFFWLVSGFFFK
1307	2657	A	10205	85	308	QGLPSTMVKLGCSFSGKPGKDPGDQDGAAM DSVPLISPLDISQLQPPLPDQVVIKTQTEYQLS SPDQQNYTKSR
1308	2658	A	10214	2	453	ECGGIRQPGPGPPPALASAPAATMNRVGGSPS AAANYLLCTNCRKVLRKDKRIRVSQPLTRGP SAFIPEKEVVQANTVDERTNFLVEEYSTSGRL DNTTQVMSLHTQYLESFLRSQFYMLRMDGPL PLPYRHYIAIMAAARHQCSYLINM
1309	2659	A	10233	45 -	421	RGWPEQQSTGRPRDVARQPRCQKEEGRRLRP RALESRTFQGSERSRWGPPLESTKENVQCGH RPAFPNSSWLPFHERLQVQNGECPWQVSIQM SRKHLCGGSILHWWWVLTAAHCFRRTLLDM AV
1310	2660	Ā	10241	243	442	AFOLFNAKCESAFLSKRNPLQRNWTVLYRRK HKKGQSAEIQKKRTRRAFKFQRAITGASLADI MAK
1311	2661	Ā	10261	751	176	LPGADYGGGHLSLRLFHLLLTSAAWVPDESQ VTLNSAICVLSTVLIMEFPDLGKHCSEKTCKQ LDFLPVKCDACKQDFCKDHFPYAAHKCPFAF QKDVHVPVCPLCNTPIPVKKGQIPDVVVGDHI DRDCDSHPGKKKEKIFTYRCSKEGCKKKEML QMVCAQCHGNFCIQHRHPLDHSCRHGSRPTI KAG
1312	2662	A	10270	3	669	STSSDEGSPSASTPMINKTGFKFSAEKPVIEVP SMTILDKKDGEQAKALFEKVRKFRAHVEDSD

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion LIYKLYVVQTVIKTAKFIFILCYTANFVNAISF EHVCKPKVEHLIGYEVFECTHNMAYMLKKL LISYISIICYYGFICLYTLFWLFRIPLKEYSFEKV REESSFSDIPDVKNDFAFLLHMVDQYDQLYS
1313	2663	A	10287	1221	266	KRFGVFLSEVSENKLREISLNHEWTFEKL GAHRVLSPAQGAQPRLRSAASVEVSMVGQR VLLLVAFLLSGVLLSEAAKILTISTLGGSHYLL LDRVSQILQEHGHNVTMLHQSGKFLIPDIKEE EKSYQVIRWFSPEDHQKRIKKHFDSYIETALD GRKESEALVKLMEIFGTQCSYLLSRKDIMDSL KNENYDLVFVEAFDFCSFLIAEKLVKPFVAIL PTTFGSLDFGLPSPLSYVPVFPSLLTDHMDFW GRVKNFLMFFSFSRSQWDMQSTFDNTIKEHF PEGSRPVLSHLLLKAELWFVNSDCAPDFARPL LPNTVYIGGLMEKPIKPVPQVSEPSAFSLGFT
1314	2664	A	10288	536	1890	NVQLAKFSSTLVFFFSCDADPSALAKYVLAL VKKDKSEKELKALCIDQLDVFLQKETQIFVEK LFDAVNTKSYLPPPEQPSSGSLKVEFFPPQEK DIKKEEITKEEREKKFSRRLNHSPPQSSSRYR ENRSRDERKKDDRSRKRDYDRNPPRRDSYRD RYNRRGRSRSYSRSRSWSKERLRERDRD RSRTRSRSRTRSRERDLVKPKYDLDRTDPLEN NYTPVSSVPSISSGHYPVPTLSSTITVIAPTHHG NNTTESWSEFHEDQVDHNSYVRPPMPKKRC RDYDEKGFCMRGDMCPFDHGSDPVVVEDVN LPGMQPFPAQPPVVEGPPPPCLPPPPPILTPPV NLRPPVPPGPLPPSLPPVTGPPPPLPPLQPSG MDAPPNSATSSVPTVVTTGIHHQPPPAPPSLFT ADTYDTDGYNPEAPSITNTSRPMYRHRVHPR AKLG
1315	2665	A	10293	447	1331	SHPLLSCPEKVSAKLRAAAEAAAEERRTRGA GSRGICAGLRSVAPGPEPLKQEEGRREWGSSI GTPSPCGSAQAAAAAAAEEATEKIPALRPALL WALLALWLCCATPAHALQCRDGYEPCVNEG MCVTYHNGTGYCKCPEGFLGEYCQHRDPCE KNRCQNGGTCVAQAMLGKATCRCASGFTGE DCQYSTSHPCFVSRPCLNGGTCHMLSRDTYE CTCQVGFTGRNPKCPGGNLNYQFNGIIVVYS GGSVPPSGTKTSKPAEHNAMGTGSKNFASGT LWVMVSGATSTSTSTL
1316	2666	A	10294	118	572	SLSMESNHKSGDGLSGTÖKEAALRALVQRTG YSLVQENGQRKYGGPPPGWDAAPPERGCEIFI GKLPRDLFEDELIPLCEKIGKIYEMRMMMDF NGNNRGYAFVTFSNKVEAKNAIKQLNNYEIR NGRLLGVCASVDNCRLFVGGIPKTKK
1317	2667	A	10301	158	1956	LLKSCGVLLSGVCIPCEGKGPTVLVIQTAVPQ DRPTKSSMRSAAKPWNPAIRAGGHGPDRVRP LPAASSGMKSSKSSTSLAFESRLSRLKRASSE DTLNKPGSTAASGVVRLKKTATAGAISELTES RLRSGTGAFTTTKRTGIPAPREFSVTVSRERSV PRGPSNPRKSVSSPTSSNTPITPTKHLRTPSTKP KQENEGGEKAALESQVRELLAEAKAKDSEIN RLRSELKKYKEKRTLNAEGTDALGPNVDGTS VSPGDTEPMIRALEEKNKNFQKELSDLEEENR VLKEKLIYLEHSPNSEGAASHTGDSSCPTSITQ ESSFGSPTGNQLSSDIDEYKKNIHGNALRTSG SSSSDVTKASLSPDASDFEHITAETPSRPLSSTS NPFKSSKCSTAGSSPNSVSELSLASLTEKIQKM EENHHSTAEELQATLQELSDQQQMVQELTAE

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion NEKLVDEKTILETSFHQHRERAEQLSQENEKL MNLLQERVKNEEPTTQEGKIELEQKCTGILE QGRFEREKLLNIQQQLTCSLRKVEEENQGAL EMIKRLKEENEKLNEFLELERHNNMMAKTL EECRVTLEGLKMENGSLKSHLQG GECFIMAAVVQQNDLVFEFASNVMEDERQL GDPAIFPAVIVEHVPGADILNSYAGLACVEEP
1010	2//0		10000	160		NDMITESSLDVAEEEIIDDDDDDITLTVEASCH DGDETIETIEAAEALLNMDSPGPMLDEKRINN NIFSSPEDDMVVAPVTHVSVTLDGIPEVMETQ QVQEKYADSPGASSPEQPKRKK
1319	2669	A	10322	169	654	MEVRMSGSVAVTRAIAVPGLLLLLIIATALSL LIGAKSLPASVVLEAFSGTCQSADCTIVLDAR LPRTLAGLLAGGALGLAGALMQTLTRNPLAD PGLLGVNAGASFAIVLGAALFGYSSAQEQLA MAFAGALVASLIVAFTGSQGGGQLSPVRLTL AGVXL
1320	2670	A	10323	441	2	KMNQVAVVIGGGQTLGAFLCHGLAAEGYRV AVVDIQSDKAANVAQEINAEYGESMAYGFG ADATSEQSVLALSRGVDEIFGRVDLLVYSAGI AKAAFISDFQLGDFDRSLQVNLVGYFLCARE FSRLMIRDGIQGRIIQINSKSDE
1321	2671	A	10332	1	453	RHRTAGPGSTISSRTDSASAPAARAMPCEYTY AKLTSDCSRPSLQWYTRAQSKMRRPRLLLKD ILKCTLLVFGVRILYILKLNYTTEECDMKNMH YVDPDHVKRAQKYAQQVLQKESPPKFAKTS MALLFEHRYSVDLLPFVQKAPTDSEA
1322	2672	A	10333	25	423	EPSNGPVVYSALGNEDDEILLLGKDIIGTFAAS ERKMRAHQVLTFLLLFVITSGASENASTSRGC GLDLLPQNVYLCDLDAIWGIVVEAVAGAGA LITLLLMLILLGRLPFIKEKEKKSPAVLHFLFL LGTLG
1323	2673	A	10334	52	426	SSLGNEDDEILSLAKDITGMFVASHRKMRAH QVLTFLLLFVITSVASENASTSRGCGLDLLPQ YVSLCDLDAIWGIVVEAAAGAGALITLLIMLI LLVRLPFFKEKEKKSPVGLHFLFLLGTLGP
1324	2674	A	10336		932	ERLCFPCMQSKIYSYMSPNKCSGMRFPLQEE NSVTHHEVKCQGKPLAGIYRKREEKRNAGN AVRSAMKSEEQKIKDARKGPLVPFPNQKSEA AEPPKTPPSSCDSTNAAIAKQALKKPIKGKQA PRKKAQGKTQQNRKLTDFYPVRRSSRKSKAE LQSEERKRIDELIESGKEEGMKIDLIDGKGRG VIATKQPSRGDFVVEYHGDLIEITDAKKREAL YAQDPSTGCYMYYFQYLSKTYCVDATRETN RLGRLINHSKCGNCQTKLHDIDGVPHLILIAS RDIAAGEELLYDYGDRSKASIEAHPWLKH
1325	2675	A	10338	3	870	PGSTISCSELKGTQCRATAGSRGRRPPMTCWL RGVTATFGRPAEWPGYLSHLCGRSAAMDLG PMRKSYRGDREAFEETHLTSLDPVKQFAAWF EEAVQCPDIGEANAMCLATCTRDGKPSARML LLKGFGKDGFRFFTNFESRKGKELDSNPFASL VFYWEPLNRQVRVEGPVKKLPEEEAECYFHS RPKSSQIGAVVSHQSSVIPDREYLRKKNEELE QLYQDQEVPKPKSWGGYVLYPQVMEFWQG QTNRLHDRIVFRRGLPTGDSPLGPMTHRGEE DWLYERLAP
1326	2676	A	10344	2	984	ARAAAHCGICRLVRWWRKRRSVMGIQTSPV LLASLGVGLVTLLGLAVGSYLVRRSRRPQVT LLDPNEKYLLRLLDKTTVSHNTKRFRFALPTA

NO. of oldered entitle entit	CEO ID	1 000 10	1 1/-4	1.000	I No. 12 a d	T 8 10 11 11	
nucleotide seq- uence USSN 09495 09496 09496 09496 09497 09495 09497 09496 09497 09496 09497 09496 09497 09496 09497 09496 09496 09497 09496 09497 09496 09497 09	SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
uence USAN verner		1	hod				
uence 99496		, - •	}				
uenoc 914 ng to first amino acid residue of peptide peptide peptide peptide sequence Physician (- V-Valine, V-M-Vincy-M-V-Vincy-M-V-Vincy-M-V-VINCY-M-V-V-V-V-V-V-V-V-V-V-V-V-V-V-V-V-V-V-	ŀ		1	1			l=Isoleucine, K=Lysine, L=Leucine,
mino seid reidue of peptide sequence printed sequence provide provide sequence pr		uence	İ	L .			
reidue of peptide poptide pept	uence			914			
Poptide sequence	}	ļ	ì	l			
	1	ŀ	ł	Į		sequence	
HHTIGLIPVEKHIYLSTRIDGSLVRPYTPVTSON EDGGYVVL NIKVYL KGVHPKPEPGEKMSQV LDSLKVQDVVERROPSQLLTYTGKGHFNIQP NKKSPPEPVAKKLGMIAOGTGITFMLQLIRA ILKVPEDPTQCFLLFANQTGKITRMLQLTRA ILKVPEDPTQCFLLFANQTGKITRMLQLTRA MKRHLPAFGDDVLVLLGSPPMVQLACHPN LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMRTTY LDKLGVSQKMVDPQNTAW SNDDSKFQGRMLEKKMSKGKGLGADQQ ATDHIKVQVKNNHLGLGATINNEDNWANQ DDFNQLAEINTCHGGETTDSSNKKKSFS LEEKSKISKNRVHYMKFTKGKDLSSRSKTDL CDFGRGRQSKKTEGGFGDSSFSTEDENETTTISAF TIGEYFAKRMALKNRPQVPVPGSDISSTQVE RXRGKRKNREATGKOVESSVLQVKAKRHTEG RYPERABAQERVAKKKSAPAEFQLRGPCWDQ SKKASQDAGDHVQPA RXRGKKRNEATGKOVESSVLQVKAKRHTEG RYPERABAQERVAKKKSAPAEFQLRGPCWDQ SKKASQDAGDHVQPA SKKASQDAGDHVQPA SKKASQDAGDHVQPA CRMAPNOCCPDCKVPGDDCPLVWQQCSHCF HMCILKWHAQQVQHCPMCRQVEWFER LPKYMTLATISABHLITTMYFFLSN LSFADICVTSTTIFKMLMNIGTQNKVTYTVAC LMVVILSGCATLLSVANDYFVAYALGPH LHYMVIMNPHLCGLLVLASWTMSALYSLQJ LMVVILSGCATLSVANSTAVYALGPH LHYMVIMNPHLCGLLVLASWTMSALYSLQJ LMVVILSGCATLSVANSTAVYALGPH LHYMVIMNPHLCGLLVLASWTMSALYSLGJ LMVVILSGCATLSVANSTAVYALGCH LHYMVIMNPHLCGLLVLASWTMSALYSLGJ LMVVILSGCATLSVANSTAVYALGPH LHYMVIMNPHLCGLLVLASWTMSALYSLGJ LMVVILSGCATLSVANSTAVYALGPH LHYMVIMNPHLCGLLVLASWTMSALYSLGJ LMVVILSGCATLSVANSTAVALYSLGJ LLVGVANSTAVALLGGGPLTGLVSNSKISSH LSTANSTAVALYVATPALDRAF LHYMVIMNPHLCGLLVLASWTMSALYSLGJ LLVGVANSTAVALLGGGPLTGLVSNSKISSH LSTANSTAVSTAVALTVATPALDRAF LHYMVINPHLCGLLVLASWTMSALAYSLGCGCCTEPPPPPDLQVQFEECKVERL FYTERAFTXATAGCGGCTEPPPPLDQVGEECKVERL FYTERAFTXATAGCGGCTEPPPPLDQVGEECKVERNET VLDQGEDLLTWSSQCKARSTTCTALASSHTTSKXSCKSTARSTATCSCGCCCTEPPPPPLDQVGEECKVERNET VLDQGEDLLTWSSQCKARSTTCTALASSHTTSKXSCKTTCTALASSHTTSKXSCKTTCTALASSHTTSKXSCKTTCTALASSHTTSKXSCKTTCTALASSTTCKSCTTCALASSTTCKSCTTCALASSTTCKSCTTCALASSTTCKSCTTCALASSTTCKTCTCALASTATCTALASTATCTATATATATATATATATATATATATA	ł	1	}	ļ	,		
EDQGYVDLVIKVYLKGVHPKFFEGGKMSQY		ļ <u> </u>	ļ		sequence	ļ	
LDSLKYQDVVERROPSQLITYTQGHPNQP	}		j]		
1327 2677 A 10345 1 968 LQSAGEGYTHYULLGYPRAYOLGHPN LDKLGYSQKMRFTY RR RYHRLSDMSMLAERRRKQKWAVDPQNTAW SNDDSKFGQRMLEKMGWSKGKGLGAQEQG ATDHIKVQVKNHLGLGATINNEDNWAHQ DDFNQLLAELNTCHGGETTDSSDKKEKKESFS LEEKSKISKNRYHYMFYTKGKDLSSSKSTDL DCIFGRRGSKKTPEGDASPSTFEENETTITSAF RYFRARMALKNRYPYPVGSDISSISTQVE RKGKKRNKEATGKDVESYLQPKAAKHTEG KPERAAGERVAKKKSAPAEEQLRGPCWDQ SSKASAQDAGDHVQPA SSKASAQDAGDHVQPA 2679 A 10351 3 964 GMAPFOCCPCKVFQDCFLWGQCSHCF HMHCILKWLHAQQVQQHCPMCRQEWKFKE GMAPFOCCPCKVFQDCFLWGQCSHCF HMHCILKWLHAQQVQQHCPMCRQEWKFKE LSYADICCTSTIFFKMIMMIQTONKYTVIACL MGMYFFLFAGESVEDQCFGCFGFF SMYLVTVLGNLLILLATISDSHLHTPMYFFLSN LSYADICCTSTIFFKMIMMIQTONKYTVIACL MGMYFFLFAGESVEDQCFGCFGFTGLF SMYLVTVLGNLLILLATISDSHLHTPMYFFLSN LSYADICCTSTIFFKMIMMIQTONKYTVIACL MGMYFFLFAGESVEDQCFGCFCDLE THYMWRHLGLILLWGTMKGQFFKKCP LHYMWINPHLGLILLWGTMKGQFFKKCP LHYMWINPHLGLILLWGTMKGQFFKKCP LHYMWINPHLGLILLWGTMKGQFFKKCP TEHGTFKFKFDSVAFGSSQEEDQCFRDLE TDPPNWQQLVSREVLLGLRPCGERQEVINEL TDPPNWQQLVSREVLLGLRPCGERQECVERV VLSSAATRNSHSSATASYMTVVTFMLNPFI EREKVKKAADHCROLLNYTNQAVEGARATCSNQ PFALEMKSSQKKDSRFQTFVQDAESNPLCRR LQLKDIPTOMQRLTKYPLLDNATYTEWPT EREKVKKAADHCROLLNYTNQAVEGARATCSNQ PFALEMKSSRQKKDSRFQTFVQDAESNPLCRR LQLKDIPTOMQRLTKYPLLDNATYTEWPT EREKVKKAADHCROLLNYTNQAVEGARATCSNQ PFALEMKSSRQKKDSRFQTFVQDAESNPLCRR LQLKDIPTOMQRLTKYPLLDNATYTEWPT EREKVKKAADHCROLLNYTNQAVEGARATCSNQ PFALEMKSSRQKALFTSNOT LGRQCFTCHT LGKYQEDGLTTUSRGCPESKLEEDGRIGSVTGL QSPRCLALARTSNOT LGRQCFTCHT LGKYQEDGLTTUSRGCPESKLEEDGRIGSVTGL QSSTCLECVARASTCHQMIMMIMTPE MPTMEFGGLDDSGGEFFDAREARISDSVTGL QSSTCLECVARSCTLLDGVPPVQUSSS STDEEVASSLTLLQRMTGHAKSQUFSCFEI STEDEVASSLTLLQRMTGHAKSQUFSCFEI STEDEVASSLTLLQRMTGHAKSQUFSCFEI STEDEVASSLTLLQRMTGHAKSQUFSCFEI STEDEVASSLTLLQRMTGHAKSQUFSCFEI STEDEVASSLTLQRMTGHAKSCACH SACAGES STEDEVASALSSLTLQRMTGHAKSCACH SACAGE				ĺ	İ		
ILK VPEDPTOCFLIFANOTEKDIILREDIELED	İ		Į.		<u> </u>	1	LDSLKVGDVVEFRGPSGLLTYTGKGHFNIQP
1327 2677 A 10345 1 968 LQSAGEGYTIFULLESPARPVQLACHEN LDKLGYSQKMEFTY 1527 2677 A 10345 1 968 LQSAGEGYTIFULLESPARPVAAVTQVQRR RYHRLSDMSMLAFRRRKQKWAVDPQNTAW SNDDSKFGQRMLEKMGWSKGKGLGAQEQG ATDHIRVQVKNNILGLGATINIDRNVLAHQ DDFNQLLAELNTCHQGETTDSSDKEEKKSTS LEEKSKISKNRVHYMKFTKGKDLSSRSKTDL DCIFGRRQSKKTPEGDASPSTFEENETTTSST TIQEYFARRMAALKNRFQVPYPGSDISETQVE RKRGKKRNKBATGKDVESYLQPKAKRHTEG KFRAEAQERVAKKSAPAEQLRGPCWDQ SSKASAQDAGDHVQPA 1328 2678 A 10346 173 439 GSAAMKVKIKCWFQVPYPGSDISETQVE RKRGKKRNKBATGKDVESYLQPKAKRHTEG CRMAPNGCCPDCKVPGDDCPLVWQQCSHCF HNHCILKWLHAQQVQQHCPMCRQEWKFRE SMYLVTVLGNLLLLATISDSHLHTTMYTFLAL LGGATGHLFTMYTFLAL LGGATGHLFTMYTHACL MQMYFFILAGFSQEFGCPFEFGLF. SMYLVTVLGNLLLLATISDSHLHTTMYTFLAL LGGATGHLFTMYTFLAL LGGATGHLFTMYTHACL MQMYFFILAGFSQEFGCPFEFGLF. LHYWMSPHLGGGPTGLILVASWTMSALYSLLQI LHWVRLSFCTALEPHFFCELNQVIQLACSDSF LHWWINFFIVALLGGGPTGLILVASWTMSALYSLLQI LHWVRLSFCTALEPHFFCELNQVIQLACSDSF LHWWINFFIVALGGGPTGLILVASWTMSALYSLLQI LHWVRLSFCTALEPHFFCELNQVIQLACSDSF LHWWINFFIVALGGGPTGLILVASWTMSALYSLLQI LHWVRLSFCTALEPHFFCELNQVIQLACSDSF LHWWINFFIVALGGGPTGLILVASWTMSALYSLLQI LHWVRLSFCTALEPHFFCELNQVIQLACSDSF LHWWINFFIVALGGGPTGLILVASWTMSALYSLLQI LHWVRLSFCTALEPHFFCELNQVIQLACSDSF LHWWINFFIVALDQWTYGWSKEISHSFE LRKIFSNLEDILQILHIGLINGMKAVKKNEST VIQUIGGELLTURSVAGGGPTGLILVASWTMSALYSLLQI LWVRLSFCTALEPHFFCELNQVIQLACSDSF LHKIFSNLEDILQILHIGLINGMKAVKKNEST VIQUIGGELLTURSVAGGGEKIKLHAAATFCSNQ FFALEMIKSRQKKDSRGTFVQDAESNPLICUL QLVDDITTAYPLLUDATYTEWPT EREKVKKAADHCRQILINYNQAVKEABNCQ RLEDYQRLDTSSLLLSEYPNVEELRIJDLIT RKMHTEGPLVWKNNRDKALFVISINSDIGAQIYE LVAQTVSERTWAGGGGEKIKHAAATFCSNQ FFALEMIKSRQKKDSRGTFVQDAESNPLICUL QSPDRULGESTILSSKPQSHSILSTGKSGSWTGL QSPDRULGESTILSSKPQSHSILSTGKSGSWTGL QSPDRULGESTILSSKPQSHSILSTGKSGSWTGL QSPDRULGESTILSSKPQSHSILSTGKSGSWTGL QSPDRULGESTILSSKPQSHSILSTGKSGSWTGL QSPDRULGESTILSSKPQSHSILSTGKSGSWTGL QSPDRULGESTILSKPQSHSVTGL QSPDRULGESTILSKPQHSPSGGGEKIKHAAHGENGGLIVAGGGSTVGL QSPDRULGESTILSKPQHSPSGGGRADSQSQUWGHSWCAGLBURGGF QDGANNLERGDVILLDGYDPVGG GSSTDEEVASSLTLQVGWGGMGSTCTGL QSPDRUHKBEADLARDGGGF GGEKIKHAUDHURGGGGEKURGGENUS GGGANNLEGGBT			l			1	NKKSPPEPRVAKKLGMIAGGTGITPMLQLIRA
MIREHLPAPGDDVLVLLCGPPMVQLACHPN			1				
LDKL_GYSQKMEFTY	İ	1	ľ	1		}	
1327 2677 A 10345 1 968 LQSAGEGYTHYLLLESPARPYAAVTQVQRA RYHRISDMSMLARRRIKQKW AVPRONTAW SNDDSKFGQRMLEKMGWSKGKGLGAQEQG ATDHIKVQVRNINILGLGATINIBDNIVLAND DDFNQLLAELNTCHGGETTDSSIKKEKKSFS LEEKSKISKNRYHYMKFTKGKDLSSRSKTLD CDIFGRQSKKTPEGDASPSTTEENETITTTSAF TIQEYFAKRMAALKNKPQVPVPGSDISSTCYC RKGGKRNKBATGKOVESVLQPKAKRITEG KPERAEAQREVAKKSAPAEEQLRGPCWDQ SSKASAQDAGDHVQPA 1328 2678 A 10346 173 439 GSAAMKVKIKCWNGVATWLWVANDENCGI CRMAPNGCCPDCKVPGDDCPLVWGQCSHCF HHHCIKWLHAQQVQHCPMCKQEWKFKE HHHCIKWLHAQQVQHCPMCKQEWKFKE HHHCIKWLHAQQVQHCPMCKQEWKFKE HHHCIKWLHAQQVQHCPMCKQEWKFKE LIYMYINMPHLCGLLVLASWTMSALYSLLQ LMVVRLSFCTALEIPHFCELNQVIQLACDSI LMVVRLSFCTALEIPHFCELNQVIQLACDSI LMVVRLSFCTALEIPHFCELNQVIQLACDSI LMVVRLSFCTALEIPHFCELNQVIQLACDSI LMVVRLSFCTALEIPHFCELNQVIQLACDSI LNHWIYFTVALLGGGFLTGIL'SYSKISSIH AISSAQGKYKAFSTCASHLSVSLFYGALDCI LMVVRLSFCTALEIPHFCELNQVIQLACDSI LHYWINGRALGHILLWGTMKQGFFKKC VLSSAATRISHSSATASSWITVYTMLNPFI TSLRNKDIKRALGHILLWGTMKQGFFKKCI TDPPNVQLVSREVLIGLKPCERRQEVINEL TDPPNVQLVSREVLIGLKPCERRQEVINEL TDPPNVQLVSREVLIGLKPCERRQEVINEL TDPPNVQLVSREVLIGLKPCERRQEVINEL TOPPNVQLVSREVLIGLKPCERRQEVINEL TVFERABYRTIKVLDQVPYQRVSREGILSPSE LRKIFSNLEDILQHIGINEQMKAVKRNETSI LQLKDIBTTQMQRLTKYPLLLDNATYTEWPT EREKVKKAADHCRQLLNYNQAVEREGILSPSE LRKIFSNLEDILQHIGINEQMKAVKRNETSI LQLKDIBTTQMQRLTKYPLLLDNATYTEWPT EREKVKKAADHCRQLLNYNQAVKEAENNQ RLEDYGRRIDTSSKLSSEYNVEGHGISVTGL QSPDRLGLGSTLISSKPQSHSLSTSGKSESCH LQLKDIBTTQMQRLTKYPLLDNATYTEWPT EREKVKKAADHCRQLLNYNQAVKEAENNQ RLEDYGRGTSSKLSSEYNVEGHGISVTGL QSPDRLGLGSTLISSKPQSHSLSTSGKSESCH LVAQTVSEKTVWQDLICMMAASVKRQSTKPI PLPQSTPGGGDNDEEDPSKLEEEQHQISVTGL QSPDRVGLGSTLISSKPQSHSLSTSGKSESCH LFVSEERWALDALRNIGLIKQLLVQQIGT LFVSEERWALDALRNIGLIKQLLVQQIGT KRGGGGDDSCASHLWMIMITE MPTMEPEGGLDDSGHFFDAREAHSDENPSE GDQAVNKEREDVNARISGNYLLLDCYDPGVG SSTDEVASSLTLQVRTIKIBAQLDHSBG GDQAVNKEREDVNIRISGADSTHQQUHSP QNTHSDGALSSFTPEFLVQQRWGAMENSCTE QSPSSCADSQSQMEVHIKEEADLBLKKVEE	1	1	{				
RYHRLSDMSMLAERRKQKWAVDPQNTAW SNDDSKFQRNLEKKIGWSKGKGLGAGEGG ATDHIKVQVKNNHLGLGATINSEDNWIAHQ DDFNQLLAELNTCHGGETTDSSDKEKEKSFS LEEKSKISKINYHYMKTKGKOLSSRSKTDL DCIFGKRQSKKTEGGDASPSTPEENETTITSAF TIQEPYAKRMALKINKQVYPYGSDISETQVE RKRGKKRNKEATGKDVESYLQPKAKRHTEG KPPRAEAQERVAKKKSAPAEEQLRGPCVDQ SSKASAQDAGDHVQPA 1328 2678 A 10346 173 439 GSAMKVIKCWNGVATWLWVANDENCGI CRMAPIGCCPCKVPGDDCLVWGQCSHCF HMECILKWLHAQQVQQHCPMCRQEWKFKE HMECILKWLHAQQVQQHCPMCRQEWKFKE AMPINCTVLGNILLILATISDSHLHTPMYFLSN LSFADICVTSTTIPKMLMNIQTQNKVITYIACL MQMYFFILFAGFENFLLSVMAYDRAVAICH LHYMVMNPHLCGLLVLASWTMSALTSLQI LMVVRLSFCTALEPHFFECLNGVQLACSDSF LNHMVIYSTVALLGGGPLTGILTSYSKISSHI AISSAQGKYKAFSTCASHLSVSLFYGALIGY VLSSAATRISHSSATASVMYTVVTPMLNPFI YSLSKATRISHSSATASVMYTVVTPMLNPFI YSLSKATRISHSSATASVMYTVVTPMLNPFI TSPPNWQQLVSRSPLLIGERGGFFKKCP TIPPPNWQQLVSRSPLLIGERGGFFKKCP TPPPNWQQLVSRSPLLIGERGGFFKKCP TPPPNWQQLVSRSVLLIGLROGFFKKCP LKIKISNLEDILQHIGLINGWKAVKKRNETSI UDQIGGDLLTWFSOPGEEKLKHAAATFCSNQ PFALEMKSRGKKDSRFOTPVGAESNPLCR REKYKKAADHCRQLINYVVQAVKBAENNQ RLEDYQRLDTSIKLSSEYNVEELRILDLIX RKMHEOPLWKVNRDKTDLYTLLLEDILV LLQKODRLVLKCHSKLLASTLADSKHTFSPVI LLSTVLVRQVATDNKALFVISMSDINGAQIYE LVAQVTSEKTVVQDLICRMASVKEQSTKPPI PLPQSTPGEGDNDEEDPSKLKEEQHOISVTGL QSFDRJCLGESTLISSRCSFRSTSTE SKSRVSVGLANGPRYNTERSTSTIS SFARRDSVGLARQOGASNPLCRR SFARRDSVGLARQOGASNPLCRR SFARRDSVGLARQOGASNPLCRR SFARRDSVGLARQOGASNPLCRR SFARRDSVGLARQOGASNPLCRR SFARRDSVGLARQOGASNPLTRITE	1207	000	 	10245		260	
SNDDSKFGQRM_EKMGWSKGKGLGAQEQG ATDHIKVQVKNNHLGLGATIDSDWAHQ DDFNQLLAELNTCHGQETTDSSDKKEKKFS LEEKSKISKNYTYHYMFTKGKDLSSSKTDL DCIFGKRQSKKTEGDASSTPEENSTTTTSAF TIQEYFAKRMAALKNIKPQVPVPGSDISETQVE RKRGKRNKEARGKDVESTV.OPKAKRHTEG KPERAEAQERVAKKKSAPAEEQLRGPCWDQ SSKASAQDAGDHVQPA 1328 2678 A 10346 173 439 GSAAMEVRIKCWNGVATWLWVANDENCGI CRMAPNGCCPDCKVPGDDCPLVWGQCSHCP HMHCILKWLHAQQVQQHCPMCRQEWKFKE SMYLVTU,GUILLILLATISDSHLHPMYFPLSN LSFADICVTSTTIPKMLMNIQTQNKVITYLACL MQVYFFILFAGFENFILLSVMAYDRVAICHP LHYMVMNPHLAGGLPLASVMSALVSLLQI LMVVRLSFCTALEPHFFCELNQVIQLACSDSF LNHRMVISTYLALGGGPLGIDYSKKISSH AISSAQGKYKAFSTCASHLSVVSLFYGALLGV YLSSAATRNSHSSATASVVYTVVTPMLNPFI TSLRNKDIKRALGIHLLWGTNKGQFFKKCP TSLRNKDIKRALGIHLLWGTNKGQFFKKCP TSLRNKDIKRALGIHLLWGTNKGQFFKKCP TFHEAFFTSVSVFGESGEDGPENDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVTRLVLDQVPY OQESEGEVERV TEHGTFKPFRKFFSVSVFGESGEDGPENDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVTRLVLDQVPY OQASNPLCRE LQLKDIPTQMQRLTKYPLLLDNIATYTEWFT EREKVKKADHCRQILNYVAQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPRVEEIRNLDLTK RKMHEGPLVWKVNRDKTDLYTILLEDILV LLQKQDDRLVRCHSKLLSEYPRVEEIRNLDLTK RKMHEGPLVWVVNRDKTDLYTILLEDILV LLQKQDDRLVRCHSKLLSEYPRVEEIRNLDLTK RKMHEGPLVWVVNRDKTDLYTILLEDILV LLQKQDDRLVRCHSKLLSEYPRVEEIRNLDLTK RKMHEGPLVWFVNRDKATDLYTILLEDILV LLQKQDDRLVRCHSKLLSEYPRVEEIRNLDLTK RKMHEGPLVWFVNRDKATDLYTILLEDILV LLQKQDDRLVRCHSKLLSEYPRVEEIRNLDLTK RKMHEGPLVWFVNRDKATDSKHFFSPVI RLSTYLVRQVATDNKALFVISMSDINGAQIVE LVVARGOFAKEQHTDGTLKEVGEDYQLAPDS HLPVSEERWALDALRNIGLIKQLLVQQLGLT EKSVQEDWHFPRYTTASQGPQTDSVIQNSE NIKAYHSGGHMFFRTGTGDIATCYSFRTSTE SFARRDSVGLARQDSQASMLVHMEMDTPE SFARRDSVGLARQDSQASMLVHMEMDTPE GDGAVNKEEKDVJURJSGNYLLLDKDYDPVQB SSTDEEVASSLTLQPMTGTDAVESTHQQQUESP QNTHISDGAMSPTHETLVQQRWGAMEVSCFEIQ OSPSCADSSQADMSYHKERGHLHKKVEE OSPSCADSSQADMSYHKERGHLHKKVEE	1327	20//	A	10345	1	968	LQSAGEGVTHVLILLESPARPVAAVTQVQRR
ATDHIKVQVKNNHLGLGATINSDNYLÅHQ DDPNQLAELTNCHGGETDSKKEKKES; LEKSKISKNRVHYMKFTKGKDLSSRSKTDL DCIFGKRQSKKTEGDASPSTFEENETTITSAF TIQFYFAKRMALKNRKQVPYGSDISETQVE RKRGKKRNKEATGKDVESYLQPKAKRHTEG KPFRAEAQERVAKKKSAPAGURGFCWDQ SSKASAQDAGDHVQPA GSAAMKVIKKCWNGVATWLWVANDENCGI CRMAFNGCCPDCKVPGDDCPLVWGQCSHCF HMHCILKWLHAQQVQQHCPMCRQEWKFKE GSAAMKVIKKCWNGVATWLWVANDENCGI CRMAFNGCCPDCKVPGDDCPLVWGQCSHCF HMHCILKWLHAQQVQQHCPMCRQEWKFKE SMYLVTVLGNLLIILATISDSHLHTPMYFFLSN LSFADICVTSTTIFKMLMMUTQNVTYTACL MOMYFFLIFAGFENFLLSWMAYDRFVALCHP LHYMVMMPHLCGILVLSWMAYDRFVALCHP LHYMVMMPHLCGILVLSWMAYDRFVALCHP LHYMVMSTTVALLGGGPLTGILTSYSKISSH ASSAQKKYKAFSTCASELSVVSLFYGALLGV YLSSAATRNSHSSATASVMYTVVTFMMSLYSLQI LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF LNHMVIYFTVALLGGGPLTGILTSYSKISSH ASSAQKKYKAFSTCASELSVVSLFYGALLGV YLSSAATRNSHSSATASVMYTVVTFMMFI YSLARNDIRKALGHILLWOTMKGQFFKKCP YLSSAATRNSHSSATASVMYTVVTFMMFI YSLARNDIRKALGHILLWOTMKGQFFKKCP THOTPKPFRKPDSVAFGESQSEDEQFENDLE TDPPNWQQLVSREVLLGILKPCIKRGEVNDE TDPNWQQLVSREVLLGILKPCIKRGEVNDE FYTERAHVRILKVLDQVFYQRVSREGILSSE LIKINSNLEDILQHIGLINGQMKAVKKRNETS VIDQIGEDLLTWFSGPGEEKI.KHAAATFCSNQ FFALEMKSRGKKDSRGFTYGVAESANGQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLIV RKMHEGPLVWKVNRDKTDLYTLLLDBILV LLQKODRLVLKCHSKLLASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDINGAQIVE LVAQTVSEKTVQDLICTMASVEGSTKPJ LPQSTPGEGDNDEEDPSKLKEEQHGISVTGI. QSPDDLGLESTLISSKPQSHLSTSGKSEVRD LFVAERGFAKEQRITOTI.KEVCOEDYQLGTT EKSVQEDWHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMFFRTGTGDIATCYSFRTTIS SFAPRDSVGLAPQDSQASNILLWOMHMIMTPE SFARDSVGLAPQDSQASNILLWOMHMIMTPE SFARDSVGLAPQDSQASNILLWOMHMIMTPE MTTMEPEGGLDDSGEHFDARRAHDDENPSG GDGAVNKEEKDVNILRISGNYLLLDGYDPVQE SSTDEEVASSLTLQPMTGITACVSTRTTE SFAPRDSVGLAPQDSQASNILLWOMHMIMTPE MTTMEPEGGLDDSGEHFDARRAHDDENPSG GDGAVNKEEKDVNILRISGNYLLLDGYDPVQE SSTDEEVASSLTLQPMTGITACVSTRTTE SFAPRDSVGLAPQDSQASNILLWOMHMIMTPE MTTMEPEGGLDDSGEHFDARRAHDDENPSG GDGAVNKEEKDVNILRISGNYLLLDGYDPVQE SSTDEEVASSLTLQPMTGITACVSTRTTE SFAPRDSVGLAPQDSGASNILLWOMHMIMTPE MTTMEPEGGLDDSGEHFDARRAHDENPSG GDGAVNKEEKDVNILRISGNYLLDGYDPVQE STDEEVASSLTLQPMTGITACVSTRTTE SFAPRDSVGLAPQUSCHENICLKYEE			1			i	
DDFNOLLAELNTCHGGETTDSSDKEKKEKSFS LEEKSKISKNRVHYMKFIKGKDLSSRSKTDL DCFGGRQSKKTPEGDASPSTPEENETITTSAF TIQEYFARRMALKNRPQVPVPGSDISETQVE RKRGKKRNKEATGKDVESYLQPKAKRHTEG KPERABAQERVAKKSAPAEEQLRGPCWDQ SSKASAQDAGDHVQPA 1328 2678 A 10346 173 439 GSAAMKVRIKCWNGYATWLWVANDENCGI CRMAFNGCCPDCKVPGDDCPLVWGQCSHCF HMHCILKWLHAQQVQHCPMCRGEWKFKE HMHCILKWLHAQQVQHCPMCRGEWKFKE MMYLVTVLGHLILLATISSHLHTPMYFFLSN LSFADICVTSTTIPKMLMNIQTQNSVITYIACL MQMYFFLIAFGFENTLSSHLHTPMYFFLSN LSFADICVTSTTIPKMLMNIQTQNSVITYIACL MQMYFFLIAFGFENTLSVSKUSSHH AISSAQGKYKAFSTCASHLSVVSLFYGALGV YLSSAATRNSHSSATASVATDRFVAICHP LHYMVIMNPHLCGILVLASWTMSALYSLLQI LMVVRLSFCTALEIPHFCELNQVIQLACSDSF LHMIVIFTVALLGGGPTFKKCP YSSAATRNSHSSATASVTVSLFYQALGV YLSSAATRNSHSSATASVTVSLFYQALGV YLSSAATRNSHSSATASVTVSFPTVOPLASSBISH AISSAQGKYKAFSTCASHLSVVSLFYGALGV YLSSAATRNSHSSATASVTVSFPTVOPLASSBISH LFFTERAFVRTLKVLDQVFQXSEGGLSFSE LRKIFSNLEDILQLHIGLNEGMKAVRKRNETS VIDQIGEDLITWFSGPGEEKLKHAAATFCSNQ PFALEMKSRQKKDSRFQTFVQDAESNPLCR LQLKDIPTQMQRLTKYPLLDNIATYTEWFT EERKVKKAADHCRQLINTVNQAVKEAENKQ RLEDYGRRLDTSSLKLSEYPNVERINLDLTX RKIMHEGPLVWKVNRDKTDLYTILLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDINGALFVISMSDNAGQIFY LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTFGEGDNDEEDPSKLKEEQHISVTGL QSPSDALGLGESTLISSKPQSISILSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDPQIAIPDS HLPVSEERWALDALRNLGILKQLLVQLGLT EKSVQEDWQHFFRYRTGGDIATCVSPRTSTE SFAPRDSVGLESTLISSKPSSISLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDPQIAIPDS NIKAYHSGEGHMPFRTGTGDIATCVSPRTSTE SFAPRDSVGLAPQDSQASNLVMDHMIMTPE MPTMETEGGLDDSGEHFFDAREABSDENSE GDQAVNKEEKDVALRTSGNYLLDGYDPVOE SSTDEEVASSLTLQPMTGTDLAVESTHQQGES GDQAVNKEEKEVANLRTSGNYLLDGYDPVOE SSTDEEVASSLTLQPMTGTGDIATCVSPRTSTE SFAPRDSVGLAPQDSQASNLVMDHMIMTPE MPTMETEGGLDDSGEHFFDAREABSDENSE GDQAVNKEEKEVANLRTSGNYLLDGYDPVOE SSTDEEVASSLTLQPMTGTGDIATCVSPRTSTE SFAPRDSVGLAPQDSQASNLVMDHMIMTPE MPTMETEGGLDDSGGEHFFDAREABSDENSE GDQAVNKEEKEVANLRTSGNYLLDGYDPVOE SSTDEEVASSLTLQPMTGTGDLATCVSPRTSTE SFAPRDSVGLAPQDSQASNLVMDHMIMTPE MPTMETEGGLDDSGQGGMETIKKEEADLEHLKKVEE	l	ì	ļ				
LEEKSKISKNRVHYMKFTIKGKDLSSRSKTDL DCIFGKROSKTTPEGDASTTTTSAF TIQEYFARMAALKNKPQVPVPGSDISETQVE RKIGKKRNKBATGKDVESYLQPKAKHTIEG KPERAEAGERVAKKKSAPAEEQLRGPCWDQ SSKASAQDAGDHYOPA GSAAMKVKIKCWNGYATWLWVANDENCGI CRMAFNGCCPDCKVPGDDCPLVWGQCSHCF HMHCLLKWLHAQQVQQHCPMCRQEWKFKE MMCLLKWLHAQQVQQHCPMCRQEWKFKE MMCPLIKVHAQQVQQHCPMCRQEWKFKE MMCPGNDTQISEFILLLGFSQPFGLQPFLFGLFL SMYLVTVLGNLLILLATISDSHLHTPMYFFLSN LSFADICVYSTTTPKMLMANTYONKYTIYIACL MQMYFFILFAGFENFLLSVMAYDRPVAICHP LHYMVINSPFLOGLILVLASWTMSALYSLLQI LMVVRLSFCTALEPHHFFCELNQVVQLACSDSF LNIHMVIYFTVALLGGGPLTGLLVSYSKIISSIH AISSAQGKYKAFSTCASHISVVSLPYGALLGV YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDIKRALGHILLWGTMKGQFFKKCP TEHGTPKPFRKFDSVAFGESQSEDECFENDLE TYPERAHVRTLKVLDQVFYQRVSREGLISPSE LARIFSNLEDILQI.HIGINEQMKAYKRNETS VIDQIGEDLITWFSGFGEEKILKHAAATFCSNQ PFALEMIKSRQKKDSRTOTTVODASEPILCRR LQKDDRITUTWFSGFGEEKILKHAAATFCSNQ PFALEMIKSRQKKDSRTOTTVODASEPILCRR LQKDDRITUTRWFSGFGEEKILKHAAATFCSNQ PFALEMIKSRQKKDSRTOTTVODASEPILCRR LQKDDRITUTRWFSGFGEEKILKHAAATFCSNQ RIEDTYGRILTBYFIDLYTLLEDILV LLQKQDRLVLRCHSKLLASTADSKHTFSPVI LSTVLVRQVATDNKALTVISMSDINGAQIVE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTRGEDDNDEEDFSKLKEEQIGISVTGL QSSPRDLGLESTILSSKPQSHISLSTSGKSEVRD LFVAERQPAKEQHTDGTLKEVGEDYQLAPDS NIKAYHSGEGMMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLESTILSSKTSTGNSEVRD HPVSEERWADDALRNLGILKQLLVQQGLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGMMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREARISDENPSE GDQAVNKEEKEQVALPISONYLLLDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQHSP QNTHSDGAISPTFELVQQRWGAMPSCFEI GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEEKERVLYALRISDSENPSE GDQAVNKEERFUNGERISLEALLEHLKKVEE			İ				ATDHIKVQVKNNHLGLGATINNEDNWLAHQ
DCFGKRQSKKTPEGDASPSTPEENTTTSAF TIQEYFAKRMALKNKPQPVPVGSDISETQVE RKRGKKRNKEATGKDVESYLQPKAKRHTEG KPERAEAQERVAKKSAPAEEQLROPCWDQ SSKASAQDAGDHVQPA SSKASAQDAGDHVQPA 1328 2678 A 10346 173 439 GSAAMKVKIKCWNGVATWLWVANDENCGI CRMAFNGCCPDCKVPGDDCLVWGQCSHCF HMHCILKWLHAQQVQHCPMCRQEWKFKE HMHCILKWLHAQQVQHCPMCRQEWKFKE MHCILKWLHAQQVQHCPMCRQEWKFKE MHCILKWLHAQQVQHCPMCRQEWKFKE AMYLVTVLGHLILLATISBHLHTPMYFFLSN LSFADICVTSTTIPKMLMNIQTQNKVITYIACL MQMYFFILFAGFENFLLSVMAYDRFVAICHP LHYMVIMNPHICGLLVLASVMSALYSLLQI LMVVRLSFCTALEPHFFCELNQVIQLACSDSF LNIMVIFFTVALLGGGPTKKCP HIMVIFTVALLGGGPTKKCP YLSRAATRNSHSSATASVMYTVVTPMLNPFI SYSKINKBIKRALGHILLWGTMKGQFFKKCP YSLRNKDIKRALGHILLWGTMKGQFFKKCP YSLRNKDIKRALGHILLWGTMKGQFFKKCP TEHGTPKPFRKFDSVAFGESQSEDQFENDLE TDPPNWQQLVSREVLLGLKFYGEKIRGEVNNEL FYTERAHVRILKVLDQVFYQRVSREGLSPSE LRKIFSNLEDIQLHIGLNEGMKAVRKNETS VIDQIGEDLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRPQTFVQDAESNPLCRR LQLKDIPTPQMQRLTKYPLLDNIATYTEWFT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYGRRLJOTSILKLSEPNYEEIRNDLTIX RKMHEGPLVWKVNRDKTIDLYTLLEDILV LLQKQDRLVLRCHSKLIASTADSKHTFSPVI KISTYLVRQVATDNKALTVISMSDINGQIYE LVAQTVSEKTVWQDLICRMAASVERGSTKPI LPVASERWALDALRNLGLIKQILLVQQIGLT EKSVQEDWQHFFRYRTASQGPQTIDSVIQNSE NIKAYHSGEGHMPFRTGTGDLATCYSPRTSTE SFARROSVGLAPQDSQASNILVMDHMMTPE MTMEFEGGLDDSGEHFFDAREAHSDENPSE GDQAVNKEEKEQVALRISGNYLLLDGYDPVVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISFTPELVQQRWGAMPSSCFEI SPSSCADSQSQMEVIHKEADLEHLKKVEE	ļ	}	ļ				DDFNQLLAELNTCHGQETTDSSDKKEKKSFS
TIQEYFAKRMAALKNIKPQVPVPGSDISETQVE RKRGKKRNKEATGKDVESYLQPKAKRHTEG KPERAEAQERVAKKKSAPAEEQLRGPCWDQ SSKASAQDAGDHYQPA SSKASAQDAGDHYQPA 1329 2679 A 10351 3 964 QMEPGNDTQISEFILLGFSQEPGLQPFLFGLFL SMYLVTVLGNLLILLATISDSHLHTMYFFLSN LSFADICVTSTTIPFMLMNGTQNN VIJACL MQMYFFILPAGFENFLSVMAYDRFVAICHP LHYMVIMNPHLCGLLVLASWTMSALYSLLQI LMYVRLSFCTALEIPHFFCELNQVIQLACSDSF LNIHMVIYFTVALLGGGPLTGLLVSYSKIISSH AISSAQGKYKAFSTCASHSLSVVSLPYGALLGV YLSSAATRNSHSSATASVMYTVVTPMLNPFI SKASAQGKYKAFSTCASHSLSVVSLPYGALLGV YLSSAATRNSHSSATASVMYTVVTPMLNPFI TSHTMANGTON VIJACL MQMYFFILPAGFENFLSVMAYDRFVAICHP LHYMVINNPHLCGLLVLASWTMSALYSLLQI LMYVRLSFCTALEIPHFFCELNQVIQLACSDSF LNIHMVIYFTVALLGGGPLTGLLVSYSKIISSH AISSAQGKYKAFSTCASHSLSVVSLPYGALLGV YLSSAATRNSHSSATASVMYTVVTPMLNPFI TSHTMANGTON VIJSAAATRNSHSSATASVMYTVVTPMLNPFI TSHTMSTON VIJSAATRNSHSSATASVMYTVVTPMLNPFI TSHTMSTON VIJSAATRNSHSSATASVMYTVVTPMLNPFI TSHTMSTON VIJSAATRNSHSSATASVMYTVVTPMLNPFI TSHTMSTON VIJSAATRNSHSSATASVMYTVVTPMLNPFI TSHTMSTON VIJSAATRNSHSSATASVMYTVVTPMLNPFI TSHTMSTON VIJSAATRNSHSSATASVMYTVVTPMLNPFI TSHTMSTON VIJSAATRNSHSSATASVMYTVVTPMLNPFI TSHTMSTON VIJSAATRNSHSSATASVMYTVVTPMLNPFI TSHTMSTON VIJSAATRNSHSSATASVMYTVVTPMLNPFI TSHTMSTON VIJSAATRNSHSSATASVMYTVVTPMLNPFI TSHTMSTON VIJSAATRNSHSSATASVMYTVVTPMLNPFI TSHTMSTON VIJSAATRNSHSSATASVMYTVVTPMLNPFI TSHTMSTON VIJSAATRNSHSSATASVMSHSSATASVMSHSSATASVMSHSATASVMSTATASVMSTATASVMSTATASVMSHSSATASVMSHSSATASVMSHSSATASVMSHSSATASVMSHSSATASVMSHSSATASVMSHSSATASVMSHSSATASVMSHSSATASVMSHSSATASVMSHSSATASVMSHSSATASVMSHSSATASVMSHSSATASVMSHSATASVMSHSSATASVMSHSATASVMSHSSATASVMSHSSATASVMSHSATASVMSHSSATASVMSHSATASVMSHSATASVMSHSSATASVMSHSSATASVMSHSAT	Ì	1					
RKRGKKRNKEATGEDVESYLQPKAKRHTEG KPERAEAQERVAKKKSAPAEEQLRGPCWDQ SSKASAQDAGDHVQPA 1328 2678 A 10346 173 439 GSAAMKVRIKCWNGVATWLWANDENCOI CRMAFINGCEPDCK VPGDDCPLVWGQCSHCF HMHCLIK WLHAQQVQQHCPMCRQEWKFKE GRAPANGCCPDCK VPGDDCPLVWGQCSHCF HMHCLIK WLHAQQVQQHCPMCRQEWKFKE MYLVTVLGNLILLIATISDSHLHITPMYFFLSN LSFADICVTSTITIEMLMNIQTONKVITYIACL MQMYFFILFAGFERFLIXISDSHLHITPMYFFLSN LSFADICVTSTITIEMLMNIQTONKVITYIACL MQMYFFILFAGFERFLIXISDSHLHITPMYFFLSN LSFADICVTSTITIEMLMNIQTONKVITYIACL MQMYFFILFAGFERFLIXIVSWAYDRFVAIGHP LHYMVIMNPHLCGLLVLASWTMSALYSLLQI LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF LNHNVIYFTVALLGGGPLTGILYSVSKISSIH AISSAQGKYKAFSTCASHLSVVSLFYGAILGV YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDIRRALGHILLWGTMKGQFFKKCP TEHGTFRFKFFSVAFFGVAFGESQSEDEQFERDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRILIVLDQVFYQRSEGGLSPSE LRKIFSNLEDILQLHIGINEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSKFQFTFVQDAESNPLCR LQLKDIPTQMQRLTKYPLLDIATYTEWPT EREKVKKAADHCRQILNYVNQAVESAENKQ RLEDYORRLDTSSLKLISFYPNVEELRNLDLTK RKMHEGPI-VWKVNRDKTDLYTLLEDILV LLQKQDDRLVLRCHSKLLASFYNVEELRNLDLTK RKMHEGPI-VWKVNRDKTDLYTLLEDILV LLQKQDDRLVLRCHSKLLASFYNVEELRNLDLTK RKMHEGPI-VWKVNRDKTDLYTLLEDILV LLQKQDDRLVLRCHSKLLASFYNVEELRNLDLTK RKMHEGPI-VWKVNRDKTDLYTLLEDILV LLQKQDDRLVLRCHSKLLASFYNVEELRNLDLTK RKMHEGPI-VWKVNRDKTDLYTLLEDILV LLQKQDDRLVLRCHSKLLASFYNVEELRNLDLTK RKMHEGPI-VWKVNRDKTDLYTLLEDILV LLQKQDDRLVLRCHSKLLASFYNVEELRNLDLTK RKMHEGPI-VWKVNRDKTDLYTLLEDILV LLQKQDDRLVLRCHSKLLASFYNVEELRNLDLTK RKMHEGPI-VWKVNRDKTDLYTLLEDILV LLQKQDDRLVLRCHSKLLASFYNVEELRNLDLTK RKMHEGPI-VWKVNRDKTDLYTLLEDILV LLQKQDDRLVLRCHSKLLASFYNVEELRNLDLTK RKMHEGPI-VWKVNRDKTDLYTLLLEDILV LLQKQDDRLVLRCHSKLLASFYNVEELRNLDLTK RKMHEGPI-VWRERWALAFQLSSHCTOR RCHTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	ļ		l				
SERABAQERVAKKKSAPAEEQLRGPCWDQ SSKASAQDAGDHVQPA			ĺ				
1328 2678 A 10346 173 439 GSAAMKVRIKCWNGVATWLWVANDENCGI CRMAFNGCCPDCKVPGDDCPLVWGQCSHCF HMHCILKWLHAQQVQQHCPMCRQEWKFKE GAPMINGCTPDCKVPGDDCPLVWGQCSHCF HMHCILKWLHAQQVQQHCPMCRQEWKFKE MYLVTVLGNLLILLATISDSHLHITPMYFFLSN LSFADICVTSTTIPKMLMNIQTONKVTTYIACL MQMYFFILAGFENFLLSVMAYDRFVAICHP LHYWIMMPHLCGLLVLASWTMSALYSILQI LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF LNHMVIYFTVALLGGGPLTGILYSYSKIISSH AISSAQGKYKAFSTCASHLSVVSLFYGAILGV YLSSAATRISHSSATASHLSVVSLFYGAILGV YLSSAATRISHSSATASHVTVVTPMLMPI YSLRINKDIKRALGIHLLWGTMKGQFFKKCP TPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRILKVLDQVGEECEVERV TEHGTFKPFRKPDSVAFGESQSEDEQFENDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRILKVLDQVGEECEVERV TEHGTFKPFRKPDSVAFGESQSEDEQFENDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRILKVLDQVAKEAENKQ RELBYQRRLDTSLKLSSTPNVEEIRNLDLTK RKMHEGPLVWKVNRDKTDLYTLLEDILV LQKQDDRLVLRCHSKLASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPL PLOSTPGGGDDDEEDSPSKLKEEQHGISVTGL QSPDRDLGLSTLISSKPQSHSLSTSGKSEVRD LFVASRGPHDPRTGTGDIATCVSPRTSTE SFAPRDSVGLAPQDSQASNILVMHMIMTPE MPTMEPEGGLDDSGEHFFDAREABSDENSE GDGAVNKEENDVILLINGSGNYLLDGYDPVQE SSTDEEVASSLTLOPMTGIPAVESTHQQHEP QNTHSDGAISPFTPEFLVQQRMEAMESYSCFEI QSPSSCADSQSQMETHFDAREABDENSES GDGAVNKEENDVILLINGSGNYLLDGYDPVQE SSTDEEVASSLTLOPMTGIPAVESTHQQHEP	ļ	}					KARGARANKEA IGKDVESYLQPKAKRHIEG
1328 2678 A 10346 173 439 GSAAMKVKIKCWNĞYATWLWVANDENCĞI CRMĀFNGCCPDCKYPGDCPLVWĞQCSHCF HMHCILKWLHAQQVQHCPMCKQEWKFKE HMHCILKWLHAQQVQHCPMCKQEWKFKE SWYLVTVLĞNİLIILATISDSHLHTPMYFFLSN LSFADICVISTITEKMLMNIQTQNKVITVIACL MQMYFFILFAGFENFLLSVMAYDRFVAICHP LYMWINPHLOGLVASWTMSALYSLLQI LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF LNHMVIYFTVALLGĞĞPLTĞİLİYSYKKISSIH AISSAQĞKYKAFSTƏSHLİSVVSKLİYĞAĞILĞVY YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSRNKDİKRALĞİHLLÜYSVİLİSSİH AISSAQĞKYKAFSTƏSHLİSVVSKLİYĞAĞILĞVY YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDİKRALĞİHLLÜYNÜKQÜĞFKKCP TEHĞTPKPFRKFDSVAFĞESQSEDEÇFENDLE TDPPNWQLVSREVLLĞİKPĞELĞÜVENEL FYTERAHVTLİKVLDQVFYQRVSREĞILSPSE LRKIFSNLEDILQİLİĞLİÇPĞEĞKÇÜLSPSE LRKIFSNLEDILQİLİĞİLĞİNĞQMKAVKRÜNETS VIDQİĞDELLİTWRŞQĞEKİLKİNAATTCSNQ PFALEMİKSRQKKDSRFQTFVQDAESNPLCR LQLKDIFTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADİNÇÜNİYVNQAVKEANKQ RLEDYQRRLDTSSLKLSEYPNVEELRILDİLV LLQKQDDRLVURCHSKILASTADSKHTSPVI KLSTVLVRQVATDNKALFYISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PIPQSTPĞEĞDNDEĞEDPSKLKEÇİĞİSVTĞL QŞPDRDLĞLESTLISSKÇĞİNİĞİSTĞKŞEVRD LFVAERÖFAKEQİTDĞİLKEVĞĞDYQLƏDİS HLPVSEERWALDALRNLĞİLKQLLVQQLĞLT EKSVQEDWQHFPRYRTASQĞQPQTDSVIQNSE NIKAYHŞĞĞĞİMPFTĞTĞIĞDIATCVSPRTSTE SFAPRDSVĞLAPQDSQASNILVMİDHIMITPE MPTMEPĞĞIDDSĞEHFFDARAHSDENNEĞ GÖDĞAVNKEKEVNYEĞİNĞĞİNYEĞL QŞSSCADSQŞQÜMEYHKIEĞADLEHLKKVEE	Ì	ĺ					
CRMAFNGCCPDCKVPGDDCPLVWGQCSHCF HMHCILKWHAQQVQQHCPMCRGEWKFKE QMEPGNDTQISEFILIGESQEPGLOPPLFGLF, SMYLVTVLGNLIILATISDSHLHTPMYFFLSN LSFADICVTSTTIPKMLMNIQTQNKVTTYIACL MQMYFILFAGFENLGSWATDSFVALCHP LHYMVIMPHLCGLLVLASWTMSALYSLLQI LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF LNHWVISTVALLGGGPLTGILYSYSKIISSH AISSAQGKYKAFSTCASHLSVVSLFYGALIGV YSLSAATRNSHSSATASVMYTVVTFMLNPFI YSLRNKDIKRALGHILLWGTMKGQFFKKCP TEHGTPKPFRKFDSVAFGESQSEDEQFENDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINSL FYTERAHVRTLKVLDQVFYQRVSREGILSPSE LRKIFSNLEDILQHIGLNEQMKAVRKNYETS VIDQIGEDLLTWFSGPGEEKLKHAAATTCSNQ PFALEMIKSRQKKDSRQTTVQDAESIPLCRR LQLKDIDTPQMQRLTKYPLLLDINIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDITK RKMHEGPLVWXVNDKTDLYTLLEDILV KLSTVLVRQVATDNKALFYNVEELRNLDITK KKMHEGPLVWXVNDKTDLYTLLEDILV LLQKQDDRLVLRCHSKLLASTADSKHTFSPVI KLSTVLVRQVATDNKALFYNVEELRNLDITK RKMHEGPLVWSVNDKTDLYTLLEDILV LLQKQDDRLVLRCHSKLLASTADSKHTFSPVI KLSTVLVRQVATDNKALFYNVEELRNLDITK RKMHEGPLVWSVNDKTDLYTLLEDILV LLVAQTVSEKTVWQDLICRMASVKEQSTKPI LFVAERQFAKEQHTDGTILKEVGEDYQLAPDS HLPVSEERWALDALRNLGILKQLLVQQLGLT EKSVQEDWQHFFRYRTASQGQTDSVIGNSE NIKAYHSGGGHMPFRTGTGDIATCVSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFRYRTASGGPQTDSVIGNSE NIKAYHSGGGHMPFRTGTGDIATCVSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMIMTPE MPTMEPEGGLDDSGEHARABHSDENPSE GDGAVNKEEKDVNILRISGNYLLLDGYDPVQE SSTIDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEILVQQRWGAMEYSCFEI QSPSSCADSQSQMEYHHKIEADLEHLKKVEE	1328	2679	_	10246	122	420	
HMHCILK WLHAQQVQHCPMCRQEWKKE 1329 2679 A 10351 3 964 QMEPGNDTQISEFLLEFSQEFGLPFLFGLFL SMYLVTVLGNLLILLATISDSHLHTPMYFFLSN LSFADICVTSTTIPKMLMNIQTQNKVITVIACL MQMYFFLFAGFENFLLSVMAYDRFVAICHP LHYMVIMNPHLCGLLVLASWTMSALYSLLQ LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF LNHWVIYFTVALLGGPPLTGILYSYSKIISSH AISSAQGKYKAFSTCASHLSVVSLFYGAILGV YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDIKRALGHILLWGTMKGGPFKKCP YSLRNKDIKRALGHILLWGTMKGGPFKKCP TEHOTPFPFFRKFDSVAFGESSCEDCEPENDLE TDPPNWQLVSREVILGILKPCEIKRQEVINEL FYTERAHVRILKVLDQVFYQRVSREGILSPSE LRKIFSNLEDILQLHIGLINGMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATTCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSSFYNVEELRNLDITK RKMHEGPLVWXVNRDKTDLYTILLEDILV LLQKQDDRLVIRCHSKKLASTADSKHTFSPVI KLSTVLVRQVATDNKALFYNIMSDNGAQIYE LVAQTVSEKTVWQDLCRMASVKEQSTKPI LVAQTVSEKTVWQDLCRMASVKEQSTKPI LVAQTVSEKTVWQLICRMASVKEQSTKPI LFVAERQFAKEQHTDGTLKEVGEDYQLAPDS HLPVSEERWALDALRNLGILKQLLVQQLGLT EKSVQEDWQHFFRYSTASQGQPTDSVIGNSE NIKAYHSGGGHMPFRTGTGDIATCVSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGILDSGEHFFDARRAHSDENTSE GDGAVNKEEKDVNLRISGNYLLLDGYDPVQE SSTIDEEVASSLTLQPMTGIPAVESTHQQQRSP QSFSSCADSQSQUREVHIKIEADLEHLKKVEE	1320	2076	Ι ^	10340	173	439	
1329 2679 A 10351 3 964 QMEPGNDTQISEFLLLGFSQEPGLQPFLGLFL SMYLVTVLGNLLIILATISDSHLHTPMYFFLSN LSFADICVTSTTIPMEMIMIOTONKVITVIACL MQMYFFLFAGFENFLLSVMAYDRFVAICHP LHYMWIMNPHLCGLLVLASWTMSALYSLLQI LMVVRLSFCTALEPHFFCELNQVIQLACSDSF LNHMVIYFTVALLGGGPLTGILYSYSKIISSH AISSAQGKYKAFSTCASHLSVVSLFYGAILGV YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDIKRALGIHLLWGTMKGQFFKKCP TEHGTPKPFRKFDSVAFGESQSEDEGFENDLE TDPPNWQQLVSREVLLGIKPCEIKRQEVINEL FYTERAHVRTLKVLDQVFYQRVSREGILSPSE LRKIFSNLEDILQLHIGLINEQMKAVRKNRISTS VIDQIGEDLLTWFSOPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQLLYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKODDRLVLRCHSKLLASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGI QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERGFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSERBWALDALRNIGII.KQLLVQQLGLT EKSVQEDWQHFPRYTTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDLATCYSPRTSTE SFAPRDSVGLAPQDSQASNILWMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLLDGYDPVQE SSTDEEVASSLTLQPMTGPAVESTHQQQHSP QNTHSDGAISPFTTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYJHKKEADLEHLKKVEE							
SMYLVTVLGNLLILLATISDSHLHTPMYFFLSN LSFADICVTSTTIPKMLMNIQTQNKVITYIACL MQMYFFILARGENFLLSVMAYDRFVAICHP LHYMVIMNPHLCGLLVLASWTMSALYSLLQI LMVVRLSFCTALEPHIFFCELNQVIQLACSDSF LNHMVIYFTVALLGGGPLTGILYSYSKIISSH AISSAQGKYKAFSTCASHLSVVSLFYGALIGV YLSSAATRNSHSSATASVMYTVVTFMLNPFI YSLRMKDIKRALGHILLWGTMKGQFFKKCP YSLRMKDIKRALGHILLWGTMKGQFFKKCP TSHKTDKRALGHILLWGTMKGQFFKKCP TEHGTPKPPRKFDSVAFGESQSEDGCPENDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL LRKIPSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ FFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIPTTOMORLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKLLASTADSKHTFSPVI KLSTYLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHISTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGILXQLLVQQLGLT EKSVQEDWGHPSYRTASSQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDLATCYSPRTSTE SFAPRDSVGLAPQDSQGASRILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLLDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYJHKKEADLEHLKKVEE	1220	2670	<u> </u>	10251	3	064	
LSFADICVTSTTIPKMLMNIQTQNKVITVIACL MQMYFFILFAGFENFLLSVMAYDRFVAICHP LHYWIMMPHLCGLLVLASWTMSALYSLLQI LMVVRLSFCTALEIPHFECELNQVIQLACSDSF LNHMVIYFTVALLGGGPLTGILYSYSKHSSH AISSAQGKYKAFSTCASHLSVVSLFYGAILGV YLSSAATRNSHSSATASVMYTVVTPMLNFFI YSLRNKDIKRALGHILLWGTMKGQFFKKCP TEHGTPKPFRKFDSVAFGESQSEDEQFENDLE TDPFNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHPKTLKVLDQVFYQNEVSEGSLSPSE LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEKLKHAAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKYKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSKLSEYPNVEEIRNLDLTK RKMIHEOPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDINGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDFSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGLLKQLLVQQLGLT EKSVQEDWMPFRTGTGDLATCYSPRTSTE SFARRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLLLDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYHKEADLEHLKKVEE	1323	2079	^	10331	3	964	QMEPGNDTQISEFLLLGFSQEPGLQPFLFGLFL
MQMYFFILFAGFENFLLSVMAYDRPVAICHP LHYMVIMNPHLCGLLVLASWTMSALYSLLQI LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF LNHMVIYFTVALLGGGPLTGILYSYSKHISSIH AISSAQGYKAFSTSASHLSVVSLFYGALIGV YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDIKRALGIHLLWGTMKGQFFKKCP 1330 2680 A 10352 34 2573 IPFLKSCCCCCLFDPPPPPLDQVQEEECEVERV TEHGTPKPFRKFDSVAFGESQSEDEOFENDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRTLKVLDQVFYQRVSREGILSPSE LRKIFSNLEDILQLHIGLINEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIPPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVURCHSKLLASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERGFAKEQHTDGTLKEVGEDYQIADDS HLPVSERWALDALRGLIKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFARRDSVGLAPQDSQASNILVMDHMIMTPE MPTIMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYHKKEADLEHLKKVEE					,		SMILVIVLGNLLILLAIISDSHLHIPMYFFLSN
LHYMVIMNPHLCGLLVLASWTMSALYSLLQI LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF LNHIMVIYFTVLGGOPLTGUL YSYSKHSSIH AISSAQGKYKAFSTCASHLSVVSLFYGAILGV YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDIKRALGHLLWGTMKGQFFKKCP YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDIKRALGHLLWGTMKGQFFKKCP YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDIKRALGHLLWGTMKGQFFKKCP TEHGTPKPFPKFDSVAFGESQSEDEQFENDLE TDPFNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRTLKVLDQVFYQRVSREGILSPSE LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEKILKHAAATFCSNQ PFALEMIKSRQKKDRFQTFVQDAESNPLCRR LQLKDIPTYQMQRLTKYPLLDDILATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEEIRNLDLTK RKMHEGPLVWKVNRDKTDLYTLLLEDILV LLQKQDDRLVLRCHSKLLSTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGILKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQMEYJHKIEADLEHLKKVEE			i '				
LMVVRLSFCTALEIPHFFCELNQVIQLACSDSF LNHMVIYFTVALLGGGPLTGILYSYSKHSSIH AISSAQGKYKAFSTCASHLSVVSLFYGAILGV YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDIKRALGIHLLWGTIMKGQFFKKCP YSLRNKDIKRALGIHLLWGTIMKGQFFKKCP TEHGTPKPFRKFDSVAFGESQSEDEGFENDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRTLKVLDQVFYQRVSREGILSPSE LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRQTFVQDAESNPLCRR LQLKDIFPTOMQRLTKYPLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVAQULICRMAASVKEQSTKPI PLPQSTPGEGDNIDEEDPSKLKEEQHGISVTGI QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERGFAKEQHTTDGTLKEVGEDYQIAPDS HLPVSEERWALDALRNLGI.LKQILVQQLGLT EKSVQEDVQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTIMEPEGGILDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTIDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYHKIEADLEILLKKVEE					•		
LNHMVIYFTVALLGGGPLTGILYSYSKHSSIH AISSAQGKYKAFSTCASHLSVVSLFYGAILGV YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDIKRALGIHLLWGTMKGQFFKKCP YSLRNKDIKRALGIHLLWGTMKGQFFKKCP YSLRNKDIKRALGIHLLWGTMKGQFFKKCP YSLRNKDIKRALGIHLLWGTMKGQFFKKCP TEHGTPKPFRKFDSVAFGESSGEDGFERDLE TDPPNVQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRTLKVLDQVFYQRVSREGILSPSE LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRKQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATIDNKALFVISMSDINGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQLAIPDS HLPVSEERWALDALRNLGIJKQULVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYHIKIEADLEHLKKVEE	Í		1				
AISSAQGKYKAFSTCASHLSVVSLFYGAILGV YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDIKRALGHILLWGTMKGQFFKKCP YSLRNKDIKRALGHILLWGTMKGQFFKKCP IPFLKSCCCCLFDPPPPLDQVQEEECEVERV TEHGTPKPFRKFDSVAFGESQSEDEQFENDLE TDPPNWQQLVSTQRVSREGILSPSE LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAJPDS HLPVSEERWALDALRNLGILKQLLVQQLGLT EKSVQEDWQHFPRYTASSQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTIDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPTPFEFLVQQRWGAMEYSCFEI QSPSSCADSQQIMEYHKIEADLEHLKKVEE					,		
YLSSAATRNSHSSATASVMYTVVTPMLNPFI YSLRNKDIKRALGIHLLWGTMKGQFFKKCP YSLRNKDIKRALGIHLLWGTMKGQFFKKCP IPILKSCCCCCLFDFPPPPLQVQEEECEVERV TEHGTPKPFRKFDSVAFGESQSEDEQFENDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRITLKVLDQVFYQRVSREGLSFSE LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIPTVMQRLTKYPLLDDNIATYTEWPT ERKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRVLRCHSKLLASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGILKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QSPSSCADSQSIMEYHKIEADLEHLKKVEE			l				
YSLRNKDIKRALGIHLLWGTMKGQFFKKCP 1330 2680 A 10352 34 2573 IPFLKSCCCCCLFDFPPPPLDQVQEEGEVERV TEHGTPKPFRKFDSVAFGESQSEDEQFENDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRTLKVLDQVFYQRVSREGILSPSE LRKIFSNLEDILQLHIGLNEQMKAVRKNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSFVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQLAIPDS HLPVSEERWALDALRNLGILKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLLDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQHSSP QSPSSCADSQSIMEYIHKIEADLEHLKKVEE							
10352 34 2573 IPFLKSCCCCLIFDFPPPLDQVQEEECEVERV TEHGTPKPFRKFDSVAFGESQSEDEQFENDLE TDPPNWQQLVSREVLLGLKPCEIKKQEVINEL FYTERAHVRTLKVLDQVFYQRVSREGILSPSE LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMDDIGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGILKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPTPPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE							
TEHGTPKPFRKFDSVAFGESQSEDEQFENDLE TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRTLKVLDQVFYQRVSREGLSPSE LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMHEGPLVWKVNRDKTIDLYTLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGILKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLLLDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE	1330	2680	A	10352	34	2573	
TDPPNWQQLVSREVLLGLKPCEIKRQEVINEL FYTERAHVRTLKVLDQVFYQRVSREGILSPSE LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAPDS HLPVSEERWALDALRNLGILKQILVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQQIMEYIHKIEADLEHLKKVEE							TEHGTPK PERKEDS VAFGES OSEDE OF ENDIE
FYTERAHVRTLKVLDQVFYQRVSREGILSPSE LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAPPDS HLPVSEERWALDALRNLGLIKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDLATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISFFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE							TDPPNWOOLVSREVI LGLKPCEIKROEVINEI
LRKIFSNLEDILQLHIGLNEQMKAVRKRNETS VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGILKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE							
VIDQIGEDLLTWFSGPGEEKLKHAAATFCSNQ PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLEDILV LLQKQDDRLVLRCHSKILASTADSKHIFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGILKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE							
PFALEMIKSRQKKDSRFQTFVQDAESNPLCRR LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGILKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLLDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE							- 1
LQLKDIIPTQMQRLTKYPLLLDNIATYTEWPT EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRI.DTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGI.I.KQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTI.QPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE					i		
EREKVKKAADHCRQILNYVNQAVKEAENKQ RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGI.LKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE					1		
RLEDYQRRLDTSSLKLSEYPNVEELRNLDLTK RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGLLKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE					1	ŀ	
RKMIHEGPLVWKVNRDKTIDLYTLLLEDILV LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGI.LKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTI.QPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE					l		
LLQKQDDRLVLRCHSKILASTADSKHTFSPVI KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGI.I.KQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE							
KLSTVLVRQVATDNKALFVISMSDNGAQIYE LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGI.I.KQLLVQQI.GLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE							
LVAQTVSEKTVWQDLICRMAASVKEQSTKPI PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGILIKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE			i		ļ		• • • • • • • • • • • • • • • • • • • •
PLPQSTPGEGDNDEEDPSKLKEEQHGISVTGL QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGLI.KQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE					i		
QSPDRDLGLESTLISSKPQSHSLSTSGKSEVRD LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGI.I.KQLI.VQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTI.QPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE							
LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS HLPVSEERWALDALRNLGI.I.KQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTI.QPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE			[1	ſ		
HLPVSEERWALDALRNLGI.LKQLLVQQLGLT EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTI.QPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE				1	ĺ	ĺ	LFVAERQFAKEQHTDGTLKEVGEDYQIAIPDS
EKSVQEDWQHFPRYRTASQGPQTDSVIQNSE NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTI.QPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE						•	
NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE SFAPRDSVGLAPQDSQASNILVMDHMIMTPE MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE	· [. 1	İ	ľ	-	
MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE				ļ		İ	NIKAYHSGEGHMPFRTGTGDIATCYSPRTSTE
MPTMEPEGGLDDSGEHFFDAREAHSDENPSE GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE			1	Ì		- 1	SFAPRDSVGLAPQDSQASNILVMDHMIMTPE
GDGAVNKEEKDVNLRISGNYLILDGYDPVQE SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE		1			1	1	MPTMEPEGGLDDSGEHFFDAREAHSDENPSE
SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE		}	1		į.	1	
QNTHSDGAISPFTPEFLVQQRWGAMEYSCFEI QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE			1	1	ļ	l	SSTDEEVASSLTLQPMTGIPAVESTHQQQHSP
QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE						I	
	Í	1	ļ	1	ſ	ĺ	QSPSSCADSQSQIMEYIHKIEADLEHLKKVEE
							SYTILCORLAGSALTDKHSDKS

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met hod	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible
1331	2681	A	10353	sequence 1	2100	nucleotide insertion AVEFAEGALTMAPWPELGDAQPNPDKYLEG AAGQQPTAPDKSKETNKTTDNTEAPVTKIELLP SYSTATLIDEPTEVDDPWNLPTLQDSGIKWSE RDTKGKILCFFQGIGRLILLLGFLYFFVCSLDIL SSAFQLVGGKMAGQFFSNSSIMSNPLLGLVIG VLVTVLVQSSSTSTSIVVSMVSSSLLTVRAAIP IIMGANIGTSITNTIVALMQVGDRSEFRRAFA GATVHDFFNWLSVLVLLPVEVATHYLEIITQL IVESFHFKNGEDAPDLLKVITKPFTKI.IVQLDK KVISQIAMNDEKAKNKSLVKIWCKTFTNKTQ INVTVPSTANCTSPSLCWTDGIQNWTMKNVT YKENIAKCQHIFVNFHLPDLAVGTILLILSLLV LCGCLIMIVKILGSVLKGQVATVIKKTINTDFP FPFAWLTGYLAILVGAGMTFIVQSSSVFTSAL TPLIGIGVITIERAYPLTLGSNIGTTTTAILAAL ASPGNALRSSLQIALCHFFFNISGILLWYPIPFT RLPIRMAKGLGNISAKYRWFAVFYLIIFFFLIP LTVFGLSLAGWRVLVGVGVVVVFIIILVLCLR LLQSRCPRVLPKKLQNWNFLPLWMRSLKPW DAVVSKFTGCFQMRCCCCCRVCCRACCLLC GCPKCCRCSKCCEDLEEAQEGQDVPVKAPET FDNITISREAQGEVPASDSKTECTAL
1332	2682	A	10354	30	1377	SQQSQPHRQGPPSLLTAPHSLDLPALPPGPR GSQGKLRRVLVPMSVKPSWGPGPSEGVTAVP TSDLGEIHNWTELLDLFNHTLSECHVELSQST KRVVLFALYLAMFVVGLVENLLVICVNWRG SGRAGLMNLYILNMAIADLGIVLSLPVWMLE VTLDYTWLWGSFSCRFTHYFYFVNMYSSIFF LVCLSVDRYVTLTSASPSWQRYQHRVRRAM CAGIWVLSAIIPLPEVVHIQLVEGPEPMCLFM APFETYSTWALAVALSTIILGFLLPFPLITVFN VLTACRLRQPGQPKSRRHCLLLCAYVAVFV MCWLPYHVTLLLITLHGTHISLHCHLVHLLY FFYDVIDCFSMLHCVINPILYNFLSPHFRGRLL NAVVHYLPKDQTKAGTCASSSSCSTQHSIIIT KGDSQPAAAAPHPEPSLSFQAHHLLPNTSPISP TQPLTPS
1333	2683	A	10358	2	884	AAGAGADGREPASERASRAEPPAVAMGQND LMGTAEDFADQFLRVTKQYLPHVARLCLIST FLEDGIRMWFQWSEQRDYIDTTWNCGYLLA SSFVFLNLLGQLTGCVLVLSRNFVQYACFGLF GIIALQTIAYSILWDLKFLMRNLALGGGLLLL LAESRSEGKSMFAGVPTMRESSPKQYMQLGG RVLLVLMFMTLLHFDASFFSIVQNIVGTALMI LVAIGFKTKLAALTLVVWLFAINVYFNAFWT IPVYKPMHDFLKYDFFQTMSVIGGLLLVVAL GPGGVSMDEKKKEW
1334	2684	A	10367	59	1562	QAWSLQVALSPFFFPASPSNSFAAAVPOLLFP ELPLPHVPGQESAKRRSARRFLIMSELTKELM ELVWGTKSSPGLSDTIFCRWTQGFVFSESEGS ALEQFEGGPCAVIAPVQAFLLKKLLFSSEKSS WRDCSQEEQKELLCHTLCDILESACCDHSGS YCLVSWLRGKTTEETASISGSPAESSCQVEHS SALAVEELGFERFHALIQKRSFRSLPELKDAV LDQYSMWGNKFGVLLFLYSVLLTKGIENIKN EIEDASEPLIDPVYGHGSQSLINLLTGHAVSN VWDGDRECSGMKLLGIHEQAAVGFLTLMEA LRYCKVGSYLKISKIPYLDCLASETHLTVFFA KDMALVAPEAPSEQARRVFQTYDPEDNGFIP DSLLEDVMKALDLVSDPEYINLMKNKLDPEG

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide]	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
cotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ĺ	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
l		ì	į	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
j		j		peptide	j	/-possible nucleotide deletion, \-possible
			<u> </u>	sequence		nucleotide insertion
			ļ			LGIILLGPFLQEFFPDQGSSGPESFTVYHYNGL
				ł	ì	KQSNYNEKVMYVEGTAVVMGFEDPMLQTD
		<u> </u>				DTPIKRCLQTKWPYIELLWTTDRSPSLN
1335	2685	Α	10375	82	2929	TRTKRRLGREKAMASPPRGWGCGELLLPFML
ł	·	Ì				LGTLCEPGSGQIRYSMPEELDKGSFVGNIAKD
				l		LGLEPQELAERGVRIVSRGRTQLFALNPRSGS
1			İ			LVTAGRIDREELCAQSPLCVVNFNILVENKM
1		[Ĭ		[KIYGVEVEIIDINDNFPRFRDEELKVKVNENA
İ		l				AAGTRLVLPFARDADVGVNSLRSYQLSSNLH
1		1	ŀ		1	FSLDVVSGTDGQKYPELVLEQPLDREKETVH
1		l	ļ		ļ	DLLLTALDGGDPVLSGTTHIRVTVLDANDNA
						PLFTPSEYSVSVPENIPVGTRLLMLTATDPDE
1			ł			GINGKLTYSFRNEEEKISETFQLDSNLGEISTL
!		ĺ	Ì		1	QSLDYEESRFYLMEVVAQDGGALVASAKVV
1	İ	l	1		}	VTVQDVNDNAPEVILTSLTSSISEDCLPGTVIA
\		İ				LFSVHDGDSGENGEIACSIPRNLPFKLEKSVD
]		1				NYYHLLTTRDLDREETSDYNITLTVMDHGTP
1		ł	·		ļ	PLSTESHIPLKVADVNDNPPNFPQASYSTSVT
1		İ		·		ENNPRGVSIFSVTAHDPDSGDNARVTYSLAE
İ		ł	l		1	DTFQGAPLSSYVSINSDTGVLYALRSFDYEQL
1.		ł		l	}	RDLQLWVTASDSGNPPLSSNVSLSLFVLDQN
			·			DNTPEILYPALPTDGSTGVELAPRSAEPGYLV
j					i	TKVVAVDKDSGQNAWLSYRLLKASEPGLFA
1		f				VGLHTGEVRTARALLDRDALKQSLVVAVED
Í		ĺ				HGQPPLSATFTVTVAVADRIPDILADLGSIKTP
1				,		IDPEDLDLTLYLVVAVAAVSCVFLAFVIVLLV
ļ		ŀ		•		LRLRRWHKSRLLQAEGSRLAGVPASHFVGV
ļ				•		DGVRAFLQTYSHEVSLTADSRKSHLIFPQPNY
ľ					'	ADTLLSEESCEKSEPLLMSDKVDANKEERRV
1						QQAPPNTDWRFSQAQRPGTSGSQNGDDTGT
Ì						WPNNQFDTEMLQAMILASASEAADGSSTLGG GAGTMGLSARYGPQFTLQHVLQGELGSDYR
ĺ			ĺ		i	QNVYIPGSNATLTNAAGKRDGKAPAGGNGN
}						KKKSGKKEKK
1336	2686	A	10379	<u> </u>	557	RPRRROPSFSCRVLVLEDPPCFRFTNSMNOEK
1330	2000	^	10377		337	LAKLQAQVRIGGKGTARRKKKVVHRTATAD
						DKKLQSSLKKLAVNNIAGIEEVNMIKDDGTVI
1						HFNNPKVQASLSANTFAITGHAEAKPITEMLP
l		1				GILSQLGADSLTSLRKLAEQFPRQVLDSKAPK
1		1				PEDIDEEDDDVPDLVENFDEASKNEAN
1337	2687	Ā	10380	1	1263	IPGSTISWSPAAARGLSVCRCCRLHPASAMDL
1		١.,	10000	•		FGDLPEPERSPRPAAGKEAOKGPLLFDDLPPA
· '						SSTDSGSGGPLLFDDLPPASSGDSGSLATSISO
!						MVKTEGKGAKRKTSEEEKNGSEELVEKKVC
1		i				KASSVIFGLKGYVAERKGEREEMQDAHVILN
(1		į i	1	DITEECRPPSSLITRVSYFAVFDGHGGIRASKF
1						AAQNI.HONLIRKFPKGDVISVEKTVKRCLLD
1						TFKHTDEEFLKQASSQKPAWKDGSTATCVLA
						VDNILYIANLGDSRAILCRYNEESOKHAALSL
		1				SKEHNPTQYEERMRIQKAGGNVRDGRVLGV
1						LEVSRSIGDGQYKRCGVTSVPDIRRCQLTPND
		l				RFILLACDGLFKVFTPEEAVNFILSCLEDEKIQ
			'		,	TREGKSAADARYEAACNRLANKAVQRGSAD
]		NVTVMVVRIGH
1338	2688	A	10385	3	589	GPSQSMAAGELEGGKPLSGLLNALAQDTFHG
					''	YPGITEELLRSQLYPEVPPEEFRPFLAKMRGIL
		Ī		!		KSIASADMDFNQLEAFLTAQTKKQGGITSDQ
1]		AAVISKFWKSHKTKIRESLMNQSRWNSGLRG
						LSWRVDGKSQSRHSAQIHTPVAIIELELGKYG
						QESEFLCLEFDEVKVNQILKTLSEVEESISTLIS

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of peptide	hod	ID NO:	beginning nucleotide	nucleotide	D-Aspartic Acid, E-Glutamic Acid,
eotide	sed-	1	in USSN	location	location corresponding	F=Phenylalanine, G=Glycine, H=Histidine, l=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	ł	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	Lance	ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine,
0000		1	1	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
			1	residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1			İ	peptide		/=possible nucleotide deletion, \=possible
				sequence	l	nucleotide insertion
						OPN
1339	2689	Α	10386	50	390	LGAMAKHHPDLIFCRKQAGVAIGRLCEKCDG
ł		l	1			KCVICDSYVRPCTLVRICDECNYGSYOGRCVI
1			1	ļ		CGGPGVSDAYYCKECTIQEKDRDGCPKIVNL
L				<u> </u>	}	GSSKTDLFYERKKYGFKKR
1340	2690	Α	10388	113	3472	SQLRKGASATHSSPSRTDCIAQMMDIYVCLK
ļ		İ	!			RPSWMVDNKRMRTASNFQWLLSTFILLYLM
1						NQVNSQKKGAPHDLKCVTNNLQVWNCSWK
[1			APSGTGRGTDYEVCIENRSRSCYQLEKTSIKIP
	1					ALSHGDYEITINSLHDFGSSTSKFTLNEQNVSL
		Ì	ļ			IPDTPEILNLSADFSTSTLYLKWNDRGSVFPHR
1	1					SNVIWEIKVLRKESMELVKLVTHNTTLNGKD
			i .			TLHHWSWASDMPLECAIHFVEIRCYIDNLHFS
	j	l	l '			GLEEWSDWSPVKNISWIPDSQTKVFPQDKVIL
						VGSDITFCCVSQEKVLSALIGHTNCPLIHLDGE
		·				NVAIKIRNISVSASSGTNVVFTTEDNIFGTVIF
			1			AGYPPDTPQQLNCETHDLKEIICSWNPGRVTA
			1			LVGPRATSYTLVESFSGKYVRLKRAEAPTNES
						YQLLFQMLPNQEIYNFTLNAHNPLGRSQSTIL
]			VNITEKVYPHTPTSFKVKDINSTAVKLSWHLP
		ĺ				GNFAKINFLCEIEIKKSNSVQEQRNVTIKGVE
		ĺ				NSSYLVALDKLNPYTLYTFRIRCSTETFWKW
]						SKWSNKKQHLTTEASPSKGPDTWREWSSDG
		İ				KNLIIYWKPLPINEANGKILSYNVSCSSDEETQ SLSEIPDPQHKAEIRLDKNDYIISVVAKNSVGS
		ł				SPPSKIASMEIPNDDLKIEQVVGMGKGILLTW
						HYDPNMTCDYVIKWCNSSRSEPCLMDWRKV
1						PSNSTETVIESDEFRPGIRYNFFLYGCRNQGY
						QLLRSMIGYIEELAPIVAPNFTVEDTSADSILV
						KWEDIPVEELRGFLRGYLFYFGKGERDTSKM
						RVLESGRSDIKVKNITDISQKTLRIADLQGKTS
						YHLVLRAYTDGGVGPEKSMYVVTKENSVGL
						IIAILIPVAVAVIVGVVTSILCYRKREWIKETFY
1						PDIPNPENCKALQFQKSVCEGSSALKTLEMNP
1		1				CTPNNVEVLETRSAFPKIEDTEIVSPVAERPEN
			!			RSDAKPENHVVESYCPPITEEEIPNPAADETGG
)						TAQVIYIDVQSMYQPQAKPEEEQENDPVGGA
,						GYKPQMHLPINSTVEDIAAEEDLDKTAGYRP
]						QANVNTWNLVSPDSPRSIDSNSEIVSFGSPCSI
(İ		NSRQFLIPPKDEDSPKSNGGGWSFTNFFQNKP
1241	2601		10300		5057	ND
1341	2691	A	10392	1	5057	MLPPKHLSATKPKKSWAPNLYELDSDLTKEP
		ļ	-	ļ		DVIIGEGPTDSEFFHQRFRNLIYVEFVGPRKTL
1				j	İ	IKLRNLCLDWLQPETRTKEEHELLVLEQYLTH
					İ	PEKLKPWVRAKKPENCEKLVTLLENYKEMY
1						QPEGESLHGVLVVSAGLRCPLGLSASTLLTW
1				Ì	ľ	SGLDNSLSWAAVGMSCVLWDIELHHDFLGV
]				- 1	i	ATKSVSTHAQGDAAQGLGGTIVRMWARDSN
}		}	į		j	LATGVLLDDNNSDVTSDDDMTRNRRESSPPH
Į .					}	SVIISFSGDRDWDRRGRSRDTEPRDRWSIITR
						NPRSRMPPRDLSLPVVAKTSFEMDREDDRDS
1			.	l	Į.	RAYESRSQDAESYQNVVDLAEDRKPHNTIQD NMENYRKLLSLGVQLAEDDGHSHMTQGHSS
		1	i 1		İ	RSKRSAYPSTSRGLKTMPEAKKSTHRRGICED
 						NANDA I FOLONGEN I MPEAKKNI MKKGICED
				l	ļ	ECCUCIONE VENETA DI ICONDO
				İ	[ESSHGVIMEKFIKDVSRSSKSGRARESSDRSQ
					{	ESSHGVIMEKFIKDVSRSSKSGRARESSDRSQ RFPRMSDDNWKDISLNKRESVIQQRVYEGNA
						ESSHGVIMEKFIKDVSRSSKSGRARESSDRSQ RFPRMSDDNWKDISLNKRESVIQQRVYEGNA FRGGFRFNSTLVSRKRVLERKRRYHFDTDGK
					<u> </u>	ESSHGVIMEKFIKDVSRSSKSGRARESSDRSQ RFPRMSDDNWKDISLNKRESVIQQRVYEGNA FRGGFRFNSTLVSRKRVLERKRRYHFDTDGK GSIHDQKGCPRKKPFECGSEMRKAMSVSSLS
						ESSHGVIMEKFIKDVSRSSKSGRARESSDRSQ RFPRMSDDNWKDISLNKRESVIQQRVYEGNA FRGGFRFNSTLVSRKRVLERKRRYHFDTDGK

SEQ ID	SEQ ID	Mct	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid, E=Glutamic Acid,
nucl-	peptide		in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine,
eotide	seq-		USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence	\	09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence			914	ng to first amino acid	acid residue of peptide	Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan.
İ				residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon,
1	}			peptide	Jaqueila	/=possible nucleotide deletion, \=possible
1				sequence		nucleotide insertion
						SQVGGKRFECKDCGETFNKSAALAEHRKIHA
ł						RGYLVECKNQECEEAFMPSPTFSELQKIYGK
						DKFYECRVCKETFLHSSALIEHQKIHFGDDKD
İ		!		'		NEREHEREREREGETFRPSPALNEFQKMYG
}	į					KEKMYECKVCGETFLHSSSLKEHQKIHTRGN
1						PFENKGKVCEETFIPGQSLKRRQKTYNKEKLC
1	İ					DFTDGRDAFMQSSELSEHQKIHSRKNLFEGR GYEKSVIHSGPFTESQKSHTITRPLESDFDEKA
1		1	ļ			FTISSNPYENQKIPTKENVYEAKSYERSVIHSL
1		1				ASVEAQKSHSVAGPSKPKVMAESTIQSFDAIN
1	ł		ł			HQRVRAGGNTSEGREYSRSVIHSLVASKPPRS
						HNGNELVESNEKGESSIYISDLNDKRQKIPAR
						ENPCEGGSKNRNYEDSVIQSVFRAKPQKSVP
1					}	GEGSGEFKKDGEFSVPSSNVREYQKARAKKK
1						YIEHRSNETSVIHSLPFGEQTFRPRGMLYECQ
ļ	1		1			ECGECFAHSSDLTEHQKIHDREKPSGSRNYE WSVIRSLAPTDPQTSYAQEQYAKEQARNKCK
1						DFRQFFATSEDLNTNQKIYDQEKSHGEESQGE
1		1				NTDGEETHSEETHGQETIEDPVIQGSDMEDPQ
1	ĺ	i I	İ	ĺ		KDDPDDKIYECEDCGLGFVDLTDLTDHQKVH
1	1					SRKCLVDSREYTHSVIHTHSISEYQRDYTGEQ
			[1		LYECPKCGESFIHSSFLFEHQRIHEQDQLYSM
	l			į .		KGCDDGFIALLPMKPRRNRAAERNPALAGSA
İ					:	IRCLLCGQGFIHSSALNEHMRLHREDDLLEQS
1	Ì					QMAEEAIIPGLALTEFQRSQTEERLFECAVCG
	ļ	}				ESFVNPAELADHVTVHKNEPYEYGSSYTHTS FLTEPLKGAIPFYECKDCGKSFIHSTVLTKHKE
	İ					LHLEEEEEDEAAAAAAAAAQEVEANVHVPQ
1				•		VVLRIQGLNVEAAEPEVEAAEPEV
			,	,		EAAEPNGEAEGPDGEAAEPIGEAGQPNGEAE
ļ	1					QPNGDADEPDGAGIEDPEERAEEPEGKAEEPE
1		ł				GDADEPDGVGIEDPEEGEDQEIQVEEPYYDC
		,				HECTETFTSSTAFSEHLKTHASMIFEPANAFG
1						ECSGYIERASTSTGGANQADEKYFKCDVCGQ
1342	2692	A	10393	2	1350	LFNDHLSLARHQNTHTG GRPRSSSDNRNFLRERAGLSSAAVQTRIGNSA
1342	2032	n l	10353	-	1330	ASRRSPAARPPVPAPPALPRGRPGTEGSTSLS
		1	.			APAVLVVAVAVVVVVVSAVAWAMANYIHV
}]] .]	j .		PPGSPEVPKLNVTVQDQEEHRCREGALSLLQ
			1	·		HLRPHWDPQEVTLQLFTDGITNKLIGCYVGN
]		1			TMEDVVLVRIYGNKTELLVDRDEEVKSFRVL
	1					QAHGCAPQLYCTFNNGLCYEFIQGEALDPKH
			l	l '		VCNPAIFRLIARQLAKIHAIHAHNGWIPKSNL
	}		1			WLKMGKYFSLIPTGFADEDINKRFLSDIPSSQI
1						LQEEMTWMKEILSNLGSPVVLCHNDLLCKNII YNEKOGDVOFIDYEYSGYNYLAYDIGNHFNE
1	[}		FAGVSDVDYSLYPDRELQSQWLRAYLEAYK
1		1				EFKGFGTEVTEKEVEILFIQVNQFALASHFFW
1	[!	GLWALIQAKYSTIEFDFLGYAIVRFNQYFKM
	L	L_	L		·	KPEVTALKVPE
1343	2693	Α	10394	102	839	PEAQTSAVLAREKGHLPTMRHEAPMQMASA
1		1	ĺ			QDARYGQKDSSDQNFDYMFKLLIIGNSSVGK
1			Į.			TSFLFRYADDSFTSAFVSTVGIDFKVKTVFKN
	1	1	[[EKRIKLQIWDTAGQERYRTITTAYYRGAMGFI
		İ		1		LMYDITNEESFNAVQDWSTQIKTYSWDNAQ
1	[1			VILVGNKCDMEDERVISTERGQHLGEQLGFE FFETSAKDNINVKQTFERLVDIICDKMSESLET
1	1	{	1	1		DPAITAAKONTRLKETPPPPQPNCAC
1344	2694	A	10395	2	4136	DRPPWNSRVDDFVTNLIHLSSKGHISPAKDTS
1		l		1		LOORTPAEMSPVLHFYVRPSGHEGAASGHTR
	·				<u> </u>	

SEQ ID NO: of nucl- eotide seq- uence	SEQ ID NO: of peptide seq- uence	Met	SEQ ID NO: in USSN 09/496 914	Predicted beginning nucleotide location correspondi ng to first amino acid residue of peptide sequence	Predicted end nucleotide location corresponding to last amino acid residue of peptide sequence	Amino acid sequence (A=Alanine C=Cysteine, D=Aspartic Acid, E=Glutamic Acid, F=Phenylalanine, G=Glycine, H=Histidine, I=Isoleucine, K=Lysine, L=Leucine, M=Methionine, N=Asparagine, P=Proline, Q=Glutamine, R=Arginine, S=Serine, T=Threonine, V=Valine, W=Tryptophan, Y=Tyrosine, X=Unknown, *=Stop codon, /=possible nucleotide deletion, \=possible nucleotide insertion RKLQGKLPELQGVETELCYNVNWTAEALPSA EETKKLMWLFGCPLLLDDVARESWLLPGSN DLLEVGPRLNFSTPTSTNIVSVCRATGLGPV DRVETTRRYRLSFAHPPSAEVEAIALATLHDR MTEQHFPHPIQSFSPESMPEPLNGPINILGEGR LALEKANQELGLALDSWDLDFYTKRFQELQR NPSTVEAFDLAQSNSEHSRHWFFKGQLHVDG QKLVHSLFESIMSTQESSNPNNVLKFCDNSSA IQGKEVRFLRPEDPTRPSRFQQQGLRHVVFT AETHNFPTGVCPFSGATTGTGGRIRDVQCTG RGAHVVAGTAGYCFGNLHIPGYNLPWEDLSF QYPGNFARPLEVAIEASNGASDYGNKFGEPV LAGFARSLGLQLPDGQRREWIKPIMFSGGIGS MEADHISKEAPEPGMEVVKVGGPVYRIGVGG GAASSVQVQGDNTSDLDFGAVQRGDPEMEQ
						KMNRVIRACVEAPKGNPICSLHDQGAGGNG NVLKELSDPAGAIIYTSRFQLGDPTLNALEIW GAEYQESNALLLRSPNRDFLTHVSARERCPA CFVGTITGDRRIVLVDDRECPVRRNGQGDAP PTPPPTPVDLELEWVLGKMPRKEFFLQRKPP MLQPLALPPGLSVHQALERVLRLPAVASKRY LTNKVDRSVGGLVAQQQCVGPLQTPLADVA VVALSHEELIGAATALGEQPVKSLLDPKVAA RLAVAEALTNLVFALVTDLRDVKCSGNWM WAAKLPGEGAALADACEAMVAVMAALGVA VDGGKDSLSMAARVGTETVRAPGSLVISAYA VCPDITATVTPDLKHPEGRGHLLYVALSPGQ HRLGGTALAQCFSQLGEHPPDLDLPENLVRA FSITQGLLKDRLLCSGHDVSDGGLVTCLLEM AFAGNCGLQVDVPVPRVDVLSVLFAEEPGLV LEVQEPDLAQVLKRYRDAGLHCLELGHTGE AGPHAMVRVSVNGAVVLEEPVGELRALWEE TSFQLDRLQAEPRCVAEEERGLRERMGPSYC LPPTFPKASVPREPGGPSPRVAILREEGSNGDR EMADAFHLAGFEVWDVTMQDLCSGAIGLDT FRGVAFVGGFSYADVLGSAKGWAAAVTFHP RAGAELRRFRKRPDTFSLGVCNGCQLLALLG WVGGDPNEDAAEMGPDSQPARPGLLLRINL SGRYESRWASVRVGPGFALMLRGMEGAVLP VWSAHGEGYVAFSSPELQAQIEARGLAPLHW ADDDGNPTEQYPLNPNGSPGGVAGICSCDGR HLAVMPHPERAVRPWQWAWRPPPFDTLTTS PWLQLFINARNWTLEGSC
1345	2695	A	10396	65	642	GVRGFWAGTMASRAGPRAAGTDGSDFQHRE RVAMHYQMSVTLKYEIKKLIYVHLVIWLLLV AKMSVGHI.RLLSHDQVAMPYQWEYPYLLSI LPSLLGLLSFPRNNISYLVLSMISMGLFSIAPLI YGSMEMFPAAQQLYRHGKAYRFLFGFSAVSI MYLVLVLAVQVHAWQLYYSKKLLDSWFTST QEKKHK
1346	2696	A	10398	1	718	DDFVRCGPQSAAMGASARLLRAVIMGAPGS GKGTVSSRITTHFELKHLSSGDLLRDNMLRGT EIGVLAKAFIDQGKLIPDDVMTRLALHELKNL TQYSWLLDGFPRTLPQAEALDRAYQIDTVINL NVPFEVIKQRLTARWIHPASGRYYNIEFNPPK TVGIDDLTGEPLIQREDDKPETVIKRLKAYED QTKPVLEYYQKKGVLETFSGTETNKIWPYVY AFLQTKVPQRSQKASVTP
1347	2697	A	10402	153	1969	KHRQENNALDMAPEHMTGPMCLIENTNGEL VANPEALKILSAITQPVVVVAIVGLYRTGKSY

SEQ ID	SEQ ID	Met	SEQ	Predicted	Predicted end	Amino acid sequence (A=Alanine C=Cysteine,
NO: of	NO: of	hod	ID NO:	beginning	nucleotide	D=Aspartic Acid. E=Glutamic Acid.
nucl-	peptide	{	in	nucleotide	location	F=Phenylalanine, G=Glycine, H=Histidine.
eotide	seq-	ļ	USSN	location	corresponding	I=Isoleucine, K=Lysine, L=Leucine,
seq-	uence		09/496	correspondi	to last amino	M=Methionine, N=Asparagine, P=Proline,
uence	ł	ł	914	ng to first	acid residue	Q=Glutamine, R=Arginine, S=Serine.
1	İ	,	i	amino acid	of peptide	T=Threonine, V=Valine, W=Tryptophan,
i i	l			residue of	sequence	Y=Tyrosine, X=Unknown, *=Stop codon.
	ļ]	<u> </u>	peptide	<u>-</u>	/=possible nucleotide deletion_\=possible
ł		j		sequence	ļ	nucleotide insertion
						LMNKLAGKNKGFSLGSTVKSHTKGIWMWCV
ļ	J	}]	J		PHPKKPEHTLVLLDTEGLGDVKKGDNONDS
	ł					WIFTLAVLLSSTLVYNSMGTINQQAMDOLYY
Į		1				VTELTHRIRSKSSPDENENEDSADFVSFFPDFV
	[['		WTLRDFSLDLEADGQPLTPDEYLEYSLKLTO
						GTSQKDKNFNLPRLCIRKFFPKKKCFVFDLPI
1		ſ				HRRKLAQLEKLQDEELDPEFVQQVADFCSYI
						FSNSKTKTLSGGIKVNGPRLESLVLTYINAISR
í I	•		Ì			GDLPCMENAVLALAQIENSAAVQKAIAHYD
1						QQMGQKVQLPAETLQELLDLHRVSEREATEV
1						YMKNSFKDVDHLFQKKLAAQLDKKRDDFCK
,						QNQEASSDRCSALLQVIFSPLEEEVKAGIYSK
1						PGGYCLFIQKLQDLEKKYYEEPRKGIQAEEIL
						QTYLKSKESVTDAILQTDQILTEKEKEIEVEC
						VKAESAQASAKMVEEMQIKYQQMMEEKEKS
])	j			YQEHVKQLTEKMERERAQLLEEQEKTLTSKL
1					i	QEQARVLKERCQGESTQLQNEIQKLQKTLKK
					ļ	KTKRYMSHKLKI
1348	2698	A	10404	5	892	TQLPAPLSGVLSRLQLGSGAPLLTWVQETAG
						VAGGAPRRRTPVTMWRLLARASAPLLRVPLS
]	·					DSWALLPASAGVKTLLPVPSFEDVSIPEKPKL
]						RFIERAPLVPKVRREPKNLSDIRGPSTEATEFT
1 1						EGNFAILALGGGYLHWGHFEMMRLTINRSM
1						DPKNMFAIWRVPAPFKPITRKSVGHRMGGGK
1			· ·			GAIDHYVTPVKAGRLVVEMGGRCEFEEVQG
				,		FLDQVAHKLPFAAKAVSRGTLEKMRKDQEE
()						RERNNQNPWTFERIATANMLGIRKVLSPYDL
1349	2699	A	10409	59	1184	THKGKYWGKFYMPKRV
1.545	2077	^	10409	75	1104	LRRNCSALGGLFQTIISDMKGSYPVWEDFINK
]				ļ		AGKLQSQLRTTVVAAAAFLDAFQKVADMAT NTRGGTREIGSALTRMCMRHRSIEAKLROFSS
		'				ALIDCLINPLOEQMEEWKKVANOLDKDHAK
		l		J		EYKKARQEIKKKSSDTLKLOKKAKKGRGDIQ
j 1			ļ			PQLDSALQDVNDKYLLLEETEKQAVRKALIE
				j		ERGRFCTFISMLRPVIEEEISMLGEITHLOTISE
		, i	.,			DLKSLTMDPHKLPSSSEQVILDLKGSDYSWS
J .						YQTPPSSPSTTMSRKSSVCSSLNSVNSSDSRSS
1 [1	. 1			GSHSHSPSSHYRYRSSNLAQQAPVRLSSVSSH
					İ	DSGFISQDAFQSKSPSPMPPEAPNQRRKEKRE
()		1	i			PDPNGGGPTTASGPPAAAEEAQRPRSM
1350	2700	A	10410	511	958	AGRGGPGKPVSWSSGPGSPGQTQRRSWVKST
1 1						RGHSSLLPPSQDFVAGLSVILRGTVDDRLNW
			1		l	AFNLYDLNKDGCITKEEMLDIMKSIYDMMG
ļ		ı		l		KYTYPALREEAPREHVESFFOKMDRNKDGV
1			Į		i	VTIEEFIESCOKDENIMRSMQLFDNVI
						· · · · · · · · · · · · · · · · · · ·

WHAT IS CLAIMED IS:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, and complementary sequences thereof.

- 2. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide hybridizes to the polynucleotide of claim 1 under stringent hybridization conditions.
- 3. An isolated polynucleotide encoding a polypeptide with biological activity, wherein said polynucleotide has greater than about 90% sequence identity with the polynucleotide of claim 1.
- 4. The polynucleotide of claim 1 wherein said polynucleotide is DNA.
- 5. An isolated polynucleotide of claim 1 wherein said polynucleotide comprises the complementary sequences.
- 6. A vector comprising the polynucleotide of claim 1.
- 7. An expression vector comprising the polynucleotide of claim 1.
- 8. A host cell genetically engineered to comprise the polynucleotide of claim 1.
- 9. A host cell genetically engineered to comprise the polynucleotide of claim 1 operatively associated with a regulatory sequence that modulates expression of the polynucleotide in the host cell.
- 10. An isolated polypeptide, wherein the polypeptide is selected from the group consisting of:
 - (a) a polypeptide encoded by any one of the polynucleotides of claim 1; and
 - (b) a polypeptide encoded by a polynucleotide hybridizing under stringent conditions with any one of SEQ ID NO:1-1350.
- 11. A composition comprising the polypeptide of claim 10 and a carrier.
- 12. An antibody directed against the polypeptide of claim 10.

13. A method for detecting the polynucleotide of claim 1 in a sample, comprising:

- a) contacting the sample with a compound that binds to and forms a complex with the polynucleotide of claim 1 for a period sufficient to form the complex; and
- b) detecting the complex, so that if a complex is detected, the polynucleotide of claim 1 is detected.
- 14. A method for detecting the polynucleotide of claim 1 in a sample, comprising:
- a) contacting the sample under stringent hybridization conditions with nucleic acid primers that anneal to the polynucleotide of claim 1 under such conditions;
- b) amplifying a product comprising at least a portion of the polynucleotide of claim 1; and
- c) detecting said product and thereby the polynucleotide of claim 1 in the sample.
- 15. The method of claim 14, wherein the polynucleotide is an RNA molecule and the method further comprises reverse transcribing an annealed RNA molecule into a cDNA polynucleotide.
- 16. A method for detecting the polypeptide of claim 10 in a sample, comprising:
- a) contacting the sample with a compound that binds to and forms a complex with the polypeptide under conditions and for a period sufficient to form the complex; and
- b) detecting formation of the complex, so that if a complex formation is detected, the polypeptide of claim 10 is detected.
- 17. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:
- a) contacting the compound with the polypeptide of claim 10 under conditions sufficient to form a polypeptide/compound complex; and
- b) detecting the complex, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 18. A method for identifying a compound that binds to the polypeptide of claim 10, comprising:

a) contacting the compound with the polypeptide of claim 10, in a cell, under conditions sufficient to form a polypeptide/compound complex, wherein the complex drives expression of a reporter gene sequence in the cell; and

- b) detecting the complex by detecting reporter gene sequence expression, so that if the polypeptide/compound complex is detected, a compound that binds to the polypeptide of claim 10 is identified.
- 19. A method of producing the polypeptide of claim 10, comprising,
- a) culturing a host cell comprising a polynucleotide sequence selected from the group consisting of a polynucleotide sequence of SEQ ID NO: 1-1350, a mature protein coding portion of SEQ ID NO: 1-1350, an active domain of SEQ ID NO: 1-1350, complementary sequences thereof and a polynucleotide sequence hybridizing under stringent conditions to SEQ ID NO: 1-1350, under conditions sufficient to express the polypeptide in said cell; and
 - b) isolating the polypeptide from the cell culture or cells of step (a).
- 20. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 1351-2700, the mature protein portion thereof, or the active domain thereof.
- 21. The polypeptide of claim 20 wherein the polypeptide is provided on a polypeptide array.
- 22. A collection of polynucleotides, wherein the collection comprises the sequence information of at least one of SEQ ID NO: 1-1350.
- 23. The collection of claim 22, wherein the collection is provided on a nucleic acid array.
- 24. The collection of claim 23, wherein the array detects full-matches to any one of the polynucleotides in the collection.
- 25. The collection of claim 23, wherein the array detects mismatches to any one of the polynucleotides in the collection.
- 26. The collection of claim 22, wherein the collection is provided in a computer-readable format.

27. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

28. A method of treatment comprising administering to a mammalian subject in need thereof a therapeutic amount of a composition comprising an antibody that specifically binds to a polypeptide of claim 10 or 20 and a pharmaceutically acceptable carrier.

Pages 340 to 1963 of this application contain amino acid sequence listings. They can be obtained at the address given below.

Los pages .340 to 1963 de cette demande contiennent des listages des séquences d'acides aminés. Elles peuvent être obtenues à l'adresse indiquée ci-dessous.

World Intellectual Property Organization 34, chemin des Colombettes CH-1211 Genève 20